

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 95 Seconds
(without alignments)
369.423 Million cell updates/sec

Title: US-09-646-679-15
Perfect score: 678
Sequence: 1 MATKSMGDIKKKLNVL.....LAKPLTKDKIIPINQLMND 136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_23:

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	678	100.0	136	10	O82445
2	538	79.4	142	10	Q9M8Y4
3	181	26.7	1969	5	O15763
4	179.5	26.5	820	16	Q8Z263
5	178.5	26.3	1197	16	Q8CVU5
6	174	25.7	394	16	Q9A3P0
7	173.5	25.6	949	16	Q8FPF9
8	173	25.5	417	2	Q9RLC7
9	171.5	25.3	933	16	Q8XE39
10	171.5	25.3	957	16	Q8ZGR4
11	171	25.2	927	2	Q9ANY0
12	170	25.1	769	16	Q8PJN8
13	167.5	24.7	1364	16	Q8PJN8
14	166.5	24.6	507	2	O85663
15	163.5	24.1	1364	16	Q8P883
16	162.5	24.0	642	16	Q9A3L2

17	162.5	24.0	719	3	Q9P896
18	162.5	24.0	778	16	Q8ZB69
19	162	23.9	767	16	Q8PD62
20	161.5	23.8	148	5	Q9GTU0
21	161.5	23.8	1709	5	Q9SPH5
22	161	23.7	716	16	Q8D5I6
23	160.5	23.7	574	16	Q8D9H9
24	159.5	23.5	571	16	Q8DAE1
25	159	23.5	117	17	Q28381
26	159	23.5	736	16	Q8KR16
27	159	23.5	1645	16	Q8YVW3
28	158.5	23.4	1266	3	Q8NIY2
29	157	23.2	900	16	Q9AC40
30	156.5	23.1	469	2	Q8RNK0
31	156	23.0	1000	17	Q8TPA1
32	155	22.9	690	16	Q9A501
33	155	22.9	757	12	Q8QKV7
34	154.5	22.8	142	17	Q28687
35	154.5	22.8	572	16	Q9KS16
36	154	22.7	708	3	Q9CIQ7
37	153.5	22.7	848	2	Q9AEW3
38	153.5	22.6	713	16	Q9A472
39	153.5	22.6	939	2	Q8KWS5
40	153	22.6	1268	2	Q8KOV0
41	152.5	22.5	120	17	Q48299
42	152.5	22.5	514	16	Q9ABT2
43	152.5	22.5	929	16	Q8BEQ2
44	152.5	22.5	935	2	Q8RNF4
45	152.5	22.5	1068	16	Q8PMW3

ALIGNMENTS

RESULT 1

O82445 PRELIMINARY; PRT; 136 AA.
AC O82445; O82445; (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Response regulator protein.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
CX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitelaw C.A., Paul W., Jenkins E.S., Taylor V.M., Roberts J.A.;
RT "A mRNA encoding a response regulator protein from Brassica napus is up-regulated during pod development."
RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF057027; AAC6225.1; -.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00072; response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50110; RESPONSE_REGULATOR; 1.
DR Phosphorylation; Sensory Transduction, KW
SQ SEQUENCE 136 AA; 15055 MW; DA99B768FAB3CF5 CRC64;

Query Match 100.0%; Score 678; DB 10; Length 136;
Best Local Similarity 100.0%; Pred. No. 4.7e-52;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKSMGDIKKKLNVLIVDDPLNLIHEKIKKAIIGISQTANNGEAVIHRDGG 60
DB 1 MATKSMGDIKKKLNVLIVDDPLNLIHEKIKKAIIGISQTANNGEAVIHRDGG 60
QY 61 SFDLLIMDKEMPERDGVSTTKKLEMEVKSMIVGVTSIADNEERRAFMEAGLNHCLAKP 120
DB 61 SFDLLIMDKEMPERDGVSTTKKLEMEVKSMIVGVTSIADNEERRAFMEAGLNHCLAKP 120

```

Qy 121 LTKDKIPIPLINQMDA 136
Db 121 LTKDKIPIPLINQMDA 136

RESULT 2
Q9MBY4 PRELIMINARY; PRT; 142 AA.
AC Q9MBY4
DT 31-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative response regulator protein (receiver component).
GN T6K12.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Rowling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome III BAC T6K12 genomic sequence.";
RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Trukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RA "Full-length messenger RNA sequences greatly improve genome
RA annotation.";
RA Genome Biol. 0:0-0(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Brover V., Trukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RA "Full-length cDNA from Arabidopsis thaliana.";
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC016829; AAF26786.1; -
DR EMBL; AY085638; AAM62859.1; -
DR InterPro; IPR001789; Response reg.
DR Pfam; PF00072; Response reg; 1.
DR ProDom; PD00039; Response reg; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
DR Phosphorylation; Sensory transduction.
KW SEQUENCE 142 AA; 15792 MW; 57487135FF2D5238 CRC64;

Query Match 79.4%; Score 538; DB 10; Length 142;
Best Local Similarity 78.9%; Pred. No. 9.8e-40;
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

Qy 1 MATKSGDIEK-----IKKKL-NVLIYDDPLNLIHEKIIKIGISQTANNGEAVII 54
Db 1 MATKSTGTEKTSIEVKKKLNVLIVDDPLNRLHEMIKTIIGISQTAKNGEAVIL 60

Qy 55 HRDGGSPDLIMKPERDGVSTTKKLEMEVKSMIVGVTSIADNEERRAFMEAGNL 114
Db 61 HRDGEAPDLIMKXENPERDGVSTTKKLEMKVTSWIVGVTSVADQEEERKAFMEAGNL 120

Qy 115 HCLAKPLTKDKIPIPLINQMDA 136
Db 121 HCLEKPLTKAKIFPLISHLEDA 142

RESULT 3
Q15763 PRELIMINARY; PRT; 1969 AA.
AC Q15763
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

Qy 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Two-component hybrid sensor and regulator.
GN LA3996.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011554; AAN51193.1; -
KW Complete proteome.
SQ SEQUENCE 820 AA; 94325 MW; 3DBE99966794567D CRC64;

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hybrid histidine kinase DHKB.
GN DHKB.
OC Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KAX3;
RC MEDLINE=98248997; PubMed=9576830;
RA Zinda M.J., Singleton C.K.;
RA "The hybrid histidine kinase dhkb regulates spore germination in
RT Dictyostelium discoideum.";
RL Dev. Biol. 196:171-183(1998).
DR EMBL; AF024654; AAB71889.1; -
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003651; His_KinA.
DR InterPro; IPR005467; His_Kinase.
DR InterPro; IPR000014; PAS domain.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; HisKA; 1.
DR Pfam; PF00072; response_reg; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00448; REC; 1.
DR TIGRFAMs; TIGR00229; sensory_box; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
DR Kinase; Phosphorylation; Sensory transduction.
KW SEQUENCE 1969 AA; 219025 MW; 8E7A7952AB1B552B CRC64;

Query Match 26.7%; Score 181; DB 5; Length 1969;
Best Local Similarity 35.8%; Pred. No. 5.1e-07;
Matches 48; Conservative 29; Mismatches 39; Indels 18; Gaps 5;

Qy 10 EIKKKLVLIYDDPLNLIHEKIIKIGISQTANNGEAVIIHRDGGSPDLIMDX 69
Db 1836 EKIEK--ILLVEDNFVNKIFSKLKDGSYIFDVAHNGVEAVCEVKKG--AYDLILMDC 1890

Qy 70 EMERDGVSTTKKLEMEVKSMI-----VGVTSIADNE--ERRAFMEAGLNHCL 117
Db 1891 QMPENDGFATTATRELEKSNLIESPPSKHSHVIVALTANSYKDKQKCLSVGMNDFL 1950

Qy 118 AKPL-TKDKIPIPLI 130
Db 1951 QRPKTSIILQMI 1964

RESULT 4
Q8EZ63 PRELIMINARY; PRT; 820 AA.
AC Q8EZ63
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Two-component hybrid sensor and regulator.
GN LA3996.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011554; AAN51193.1; -
KW Complete proteome.
SQ SEQUENCE 820 AA; 94325 MW; 3DBE99966794567D CRC64;

```

[illegible]

RT of uropathogenic Escherichia coli. ;
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016763; AN81215.1; --
KW Transferase; Complete proteome.
SQ SEQUENCE 949 AA; 106590 MW; DF8CA47F9EEB4088 CRC64;

Query Match 25.6%; Score 173.5; DB 16; Length 949;
Best Local Similarity 33.6%; Pred. No. 9.7e-07;
Matches 37; Conservative 35; Mismatches 33; Indels 5; Gaps 3;

QY 18 VLIVDDPPLNLIHEKIIKAIGISQTANNGEEAVIHRDGGSSFDLILMDKEMPERDG 76
Db 827 ILVDDHPINRLLADQLGSLGYQCKTANDGVDAALNVLK---NHIDIVLSDVNMVNDG 883
QY 77 VSTTKLREMEVKSMTIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKI 126
Db 884 YRLTQIRQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDVI 932

RESULT 8
Q9RLC7 ID Q8X39 PRELIMINARY; PRT; 417 AA.
AC Q9RLC7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative histidine kinase (fragment).
GN GACS.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JN300;
RA Graupner S., Wackernagel W.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ249741; CAB56474.1; --
DR InterPro; IPR002570; Hpt.
DR InterPro; IPR001789; Response reg.
DR Pfam; PF00072; response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00073; HPT; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01110; RESPONSE REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction.
FT NON TER 1
SQ SEQUENCE 417 AA; 45749 MW; D196B4FFDDA23BB2 CRC64;

Query Match 25.8%; Score 173; DB 2; Length 417;
Best Local Similarity 32.8%; Pred. No. 4e-07;
Matches 39; Conservative 35; Mismatches 37; Indels 8; Gaps 3;

QY 17 NVLIVDDPPLNLIHEKIIKAIGISQTANNGEEAVIHRDGGSSFDLILMDKEMPERDG 76
Db 169 SVLVCDDNPANLMLLETLTMDGGEVAVSSGQALEVVQQ--QSFDMVFDVQNGMDG 226

QY 77 VSTTKLREMEVKSMTIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKIPLI 130
Db 227 RQTEAIRRELESGGPPPLFVALTAFA-LSNERRSLQSLDYLTKPISEBQLAQV 284

RESULT 9
Q8XE39 ID Q8XE39 PRELIMINARY; PRT; 933 AA.
AC Q8XE39;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Sensor for ctr capsule biosynthesis, probable histidine kinase.
DE acting on RC8B.
GN RCSC OR Z3477 OR ECS3107.
OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIND 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22 (2001).
DR EMBL; AE005454; BAG57353.1; --
DR EMBL; AF002560; BAG36530.1; --
DR HSSP; P06143; ID4Z.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003661; His_kinA.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR000014; PAS_domain.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_C; 1.
DR Pfam; PF00512; HisKA; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HisKA; 1.
DR SMART; SM00091; PAS; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS01109; HIS_KIN; 1.
DR PROSITE; PS01110; RESPONSE REGULATORY; 1.
KW Kinase; Complete proteome.
SQ SEQUENCE 933 AA; 104605 MW; 5755C05F713E561D CRC64;

Query Match 25.3%; Score 171.5; DB 16; Length 933;
Best Local Similarity 33.9%; Pred. No. 1.4e-06;
Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;

QY 18 VLIVDDPPLNLIHEKIIKAIGISQTANNGEEAVIHRDGGSSFDLILMDKEMPERDG 77
Db 811 ILVDDHPINRLLADQLGSLGYQCKTANDGVDAALNVL--SKNHIDIVLSDVNMVNDG 868

QY 78 STTKLREMEVKSMTIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKI 126
Db 869 RLTIQIRQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDVI 916

RESULT 10
Q8ZGR4 ID Q8ZGR4 PRELIMINARY; PRT; 957 AA.
AC Q8ZGR4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Two component sensor kinase/response regulator protein Rcsc
DE (EC 2.7.3.-) (Sensor for ctr capsule biosynthesis).
GN RCSC OR YP01217 OR Y2971.
OS Yersinia pestis.

```

JC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
JC Enterobacteriaceae; Yersinia.
JC NCBi_TaxID=632;
JC [1]
JC SEQUENCE FROM N.A.
JC STRAIN=CO-92 / Biovar Orientalis;
JC MEDLINE=21470413; PubMed=11586360;
JC A Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
JC A Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
JC A Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Rarraga A.M.,
JC A Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
JC A Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
JC A Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
JC A Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
JC "Genome sequence of Yersinia pestis, the causative agent of plague."
JC Nature 413:523-527(2001).
JC [2]
JC SEQUENCE FROM N.A.
JC STRAIN=KIM5 / Biovar Mediaevalis;
JC MEDLINE=22137863; PubMed=12142430;
JC A Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Lies P.,
JC A Peria N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
JC A Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
JC A Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
JC A Perry R.D.;
JC "Genome sequence of Yersinia pestis KIM."
JC J. Bacteriol. 184:4601-4611(2002).
JC EMBL; AJ414147; CAC90055.1; -.
JC EMBL; AE013898; AAM86522.1; -.
JC HSP; P06143; I04Z.
JC InterPro; IPR003594; ATPbind_ATPase.
JC InterPro; IPR004358; Bact_sens_pr_C.
JC InterPro; IPR003661; His_kinase.
JC InterPro; IPR005467; His_kinase.
JC InterPro; IPR001789; Response_reg.
JC Pfam; PF02518; HATPase_C; 1.
JC Pfam; PF00512; HsKa; 1.
JC Pfam; PF00072; response_reg; 1.
JC PRINTS; PR00344; BCTRLSENSOR.
JC ProDom; PD000039; Response_reg; 1.
JC PROSITE; PS50109; HIS_KIN; 1.
JC PROSITE; PS50110; RESPONSE_REGULATORY; 1.
JC Kinase; Transferase; Complete proteome.
JC SEQUENCE 957 AA; 108591 MW; 0447A11F59100011 CRC64;

Query Match 25.3%; Score 171.5; DB 16; Length 957;
Best Local Similarity 35.1%; Pred. No. 1.5e-06;
Matches 39; Conservative 29; Mismatches 40; Indels 3; Gaps 2;

QY 16 LNVLIYDDPLNLIHEKIIKAGISQTNNGEEAVIIHRDGGSSFDLILMDKEMPERD 75
DB 835 LQILVVDHPINRLLADQLTTLGVRVITANDGLDVALNT--NTVDVLTVDVNPMD 892

QY 76 GVSTTKKLRMEVKSIMVGVTSIADNBEERRAPFMEAGLNHCLAKPLTKDKI 126
DB 893 GYRLTERLRQLNHNFPPIIGVTANALAEKQRC-IEAGMDNCLSKPVTDLTL 942

RESULT 11
ID Q9ANYO PRELIMINARY; PRT; 927 AA.
AC Q9ANYO;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Two-component system sensor protein.
DE RSCS.
DE RSCS.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBi_TaxID=668;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21142508; PubMed=11208780;
RX Visick K.L., Skoufos L.M.;
RX "A two-component sensor required for normal symbiotic colonization of
RX Euprymna scolopes by Vibrio fischeri."
RX J. Bacteriol. 183:835-842(2001).
RX CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
RX KINASES.
RX EMBL; AE319618; AAG60694.1; -.
RX InterPro; IPR003594; ATPbind_ATPase.
RX InterPro; IPR004358; Bact_sens_pr_C.
RX InterPro; IPR003661; His_kinase.
RX InterPro; IPR005467; His_kinase.
RX InterPro; IPR002570; Hpt.
RX InterPro; IPR001610; PAC.
RX InterPro; IPR000700; PAS-assoc_C.
RX InterPro; IPR000044; PAS_domain.
RX InterPro; IPR001789; Response_reg.
RX Pfam; PF02518; HATPase_C; 1.
RX Pfam; PF00512; HsKa; 1.
RX Pfam; PF01627; Hpt; 1.
RX Pfam; PF00785; PAC; 1.
RX Pfam; PF00989; PAS; 1.
RX Pfam; PF00072; response_reg; 1.
RX PRINTS; PR00344; BCTRLSENSOR.
RX ProDom; PD000039; Response_reg; 1.
RX SMART; SM00387; HATPase_C; 1.
RX SMART; SM00388; HsKa; 1.
RX SMART; SM00073; HPT; 1.
RX SMART; SM00086; PAC; 1.
RX SMART; SM00091; PAS; 1.
RX SMART; SM00448; REC; 1.
RX TIGRFAMs; TIGR00229; sensory_box; 1.
RX PROSITE; PS50109; HIS_KIN; 1.
RX PROSITE; PS50113; PAC; 1.
RX PROSITE; PS50112; PAS; 1.
RX PROSITE; PS50110; RESPONSE_REGULATORY; 1.
RX Kinase; Phosphorylation; Sensory transduction; Transferase.
RX SEQUENCE 927 AA; 105590 MW; 9EE0659CB2E8BF9B CRC64;

Query Match 25.2%; Score 171; DB 2; Length 927;
Best Local Similarity 33.9%; Pred. No. 1.6e-06;
Matches 43; Conservative 28; Mismatches 42; Indels 14; Gaps 4;

QY 16 LNVLIYDDPLNLIHEKIIKAGISQTNNGEEAVIIHRDGGSSFDLILMDKEMPERD 75
DB 659 LSLIADNKNALVAQMFQCRIGHKAIIEAGK--VAIDKUKETHFDLIINDNEMPMD 716

QY 76 GVSTTKKLRMEVKSIMVGVTSIADNBEERRAPFMEAGLNHCLAKPLTKDKI 129
DB 717 GILATKIIEKLIKISTVIFAYTANAFQKADN-----FLKAGANYVLTKEQENDFICA 770

QY 130 INQLMDA 136
DB 771 IKQYQDA 777

RESULT 12
ID Q8FQ37 PRELIMINARY; PRT; 769 AA.
AC Q8FQ37;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Two-component system sensor protein.
DE XAC0494.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBi_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RC MEDLINE=22022145; PubMed=12024217;
RX

```

RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,	
RA	Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,	
RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.F., Moon D.H.,	
RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,	
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,	
RA	Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,	
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,	
RA	Setubal J.C., Kitajima J.P.	
RA	Comparison of the genomes of two Xanthomonas pathogens with differing	
RT	host specificities."	
RRL	Nature 417:459-463(2002).	
RRL	EMBL; AE011887; AAM37343.1; --	
DDR	InterPro; IPR003594; ATPbind ATPase.	
DDR	InterPro; IPR004358; Bact_sens_pr_C.	
DDR	InterPro; IPR006189; CHASE.	
DDR	InterPro; IPR003661; His_kinA.	
DDR	InterPro; IPR005467; His_kinase.	
DDR	InterPro; IPR002570; Hpt.	
DDR	InterPro; IPR016110; PAC.	
DDR	InterPro; IPR000700; PAS-assoc_C.	
DDR	InterPro; IPR000014; PAS_domain.	
DDR	InterPro; IPR001789; Response_reg.	
DDR	Pfam; PF03924; CHASE; 1.	
DDR	Pfam; PF02518; HATPase_c; 1.	
DDR	Pfam; PF00512; Hiska; 1.	
DDR	Pfam; PF01627; Hpt; 1.	
DDR	Pfam; PF00785; PAC; 2.	
DDR	Pfam; PF00989; PAS; 2.	
DDR	Pfam; PF00072; response_reg; 2.	
DDR	PRINTS; PR00344; BCTRLSENSOR.	
DDR	ProDom; PD000039; Response_reg; 2.	
DDR	TIGRFAMs; TIGR00229; sensory_box; 2.	
DDR	PROSITE; PS50839; CHASE; 1.	
DDR	PROSITE; PS50109; HIS_KIN; 1.	
DDR	PROSITE; PS50113; PAC; 2.	
DDR	PROSITE; PS50112; PAS; 2.	
DDR	PROSITE; PS50110; RESPONSE_REGULATORY; 2.	
DDR	Complete proteome.	
SK	SEQUENCE 1364 AA; 150282 MW; F7CCA3856B7E7165 CRC64;	
QY	Query Match	24.7%; Score 167.5; DB 16; Length 1364;
DB	Best Local Similarity	29.2%; Pred. No. 5e-06;
QY	Matches 35; Conservative 37; Mismatches 45; Indels 3; Gaps 2;	
DB	17 NVLIVDDDPNLNLIHKKIKATGGISQTANNCBAAVITHRDGSGFDLLIMDKEMPERDG 76	
DB	1033 HLLVDDSDINCEVQRILEGEGAMVTVAHDGEQAVSTLKRAPNLFHLVLMVDQVPPVVG 1092	
QY	77 VSTTKLREMEVKSM--IVGVTSLADNEERAFMEAGLNCHLCAKPLPKDKIIINQLM 134	
DB	1093 YEATRLRQIPALASLPVIALTAGAPFPQEXA-LEAGMNGFIKAFPNVESLVAIRHFL 1151	
RESULT 14		
085663	PRELIMINARY; PRT; 507 AA.	
ID	085663	
AC	085663;	
DT	01-NOV-1998 (TrEMBLrel. 08, Created)	
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)	
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
DE	RCsc (Fragment).	
GN	RCSC.	
OS	Proteus mirabilis.	
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
OC	Enterobacteriaceae; Proteus.	
OX	NCBI_TaxID=594;	
OX	[1]_TaxID=594;	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BB2000;	
RX	MEDLINE=982047557; PubMed=9829920;	
RA	Belas R., Schneider R., Meich M.;	
RT	"Characterization of Proteus mirabilis precocious swarming mutants:	
RT	Identification of rtsA, encoding a regulator of swarming behavior."	

```

RL J. Bacteriol. 180:6126-6139(1998).
CC -|- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AF071215; AAC82662.1; -.
DR HSSP: P06857; 2CHF.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR003661; His_kinA.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR Pfam; PF00072; response_reg; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Kinase; Phosphorylation; Sensory transduction; Transferase.
FT NON_TER 1
SQ SEQUENCE 507 AA; 57326 MW; F5DA2EAD9C35DEFS CRC64;

Query Match 24.6%; Score 166.5; DB 2; Length 507;
Best Local Similarity 36.4%; Pred. No. 1.9e-06;
Matches 39; Conservative 26; Mismatches 39; Indels 3; Gaps 2;

QY 16 LNVLIYDDPDLNLIHKEIKIIGISOTANNGEAVIHRDGGSSFDLIIMDKEMPERD 75
DB 397 LTVLIYDDHPINRLTLQKKIGTATAEADGCDALAFMOE--NHVDIILTVNFMN 454
QY 76 GVSTTKKIREMVKSMIVGVTSIADNEERRAFNAGLNHCLAKPLT 122
DB 455 GIQLATTVELSSTPIIGVTANATAEKKQRC-IDAGMNDVCVRPVS 500

RESULT 15
Q8P883 PRELIMINARY; PRT; 1364 AA.
AC Q8P883
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Two-component system sensor protein.
GN XCC2360.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Parah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Gicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Kocai E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE012344; AAM41638.1; -.

```

```

DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR006189; CHASE.
DR InterPro; IPR003661; His_kinA.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR002570; Hpt_kinase.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000700; PAS-assoc_C.
DR InterPro; IPR000014; PAS_domain.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF03924; CHASE; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR Pfam; PF01627; Hpt; 1.
DR Pfam; PF00785; PAC; 2.
DR Pfam; PF00989; PAS; 2.
DR Pfam; PF00072; response_reg; 2.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS50839; CHASE; 1.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50113; PAC; 2.
DR PROSITE; PS50112; PAS; 2.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 2.
KW Complete proteome.
SQ SEQUENCE 1364 AA; 150167 MW; 48F1C5FBD9710316 CRC64;

Query Match 24.1%; Score 163.5; DB 16; Length 1364;
Best Local Similarity 30.3%; Pred. No. 1.1e-05;
Matches 36; Conservative 36; Mismatches 44; Indels 3; Gaps 3;

QY 18 LVLIYDDPDLNLIHKEIKIIGISOTANNGEAVIHRDGGSSFDLIIMDKEMPERDGV 77
DB 1034 LLLVDDSEINCEVAQRILEGEGAMVTVAHDGEQAVNTLKRAPDLFHLVLMVDQMPVVDGY 1093
QY 78 STTKKLREM-EVKSM-IVGVTSIADNEERRAFNAGLNHCLAKPLTKKIPLINQLM 134
DB 1094 EATRRLRQIPSLASLPVIALTAGAFRQCKA-LEAGMNGFIAPKPNVELVTAIRHFL 1151

Search completed: August 14, 2003, 16:49:03
JOB time : 98 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

JM protein - protein search, using sw model

run on: August 14, 2003, 16:46:47 ; Search time 23 Seconds
(without alignments)

278.071 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678

Sequence: 1 MATKMGDIKIKKLNVL.....LAKPLTKXIIPLINQLMDA 136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	178.5	26.3	1197	1 EVGS_ECOLI	P30855 escherichia
2	176.5	26.0	1197	1 EVGS_ECOLI	P58402 escherichia
3	173.5	25.6	948	1 RCSC_SALTI	Q56138 salmonella
4	172.5	25.4	948	1 RCSC_SALTY	P58662 salmonella
5	171.5	25.3	949	1 RCSC_ECOLI	P14376 escherichia
6	156.5	23.1	777	1 PRZE_WYXA	P18759 myxococcus
7	150	22.1	622	1 SKN1_YEAST	P38889 saccharomyc
8	149.5	22.1	778	1 ARCB_ECOLI	P58383 escherichia
9	149.5	22.1	778	1 ARCB_ECOLI	P22763 escherichia
10	145	21.4	1220	1 SLN1_YEAST	P39928 saccharomyc
11	143.5	21.2	712	1 SSK1_YEAST	Q07084 saccharomyc
12	142	20.9	907	1 GACS_PSESY	P48027 pseudomonas
13	140.5	20.7	226	1 DCTR_BACSU	P56602 bacillus su
14	140.5	20.7	230	1 DCTR_BACHD	Q9K998 bacillus ha
15	134	19.8	119	1 CHBY_BACSU	P44072 bacillus su
16	134	19.8	1238	1 EVGS_BORPE	P16575 bordetella
17	133.5	19.7	228	1 DCTR_BACME	P39486 bacillus me
18	133.5	19.7	918	1 BARA_ECOLI	P26607 escherichia
19	132.5	19.5	131	1 YSO1_PLEBO	P51586 plectonema
20	132.5	19.5	918	1 BAZA_SHIFL	P59342 shigella fl
21	132	19.5	859	1 LUXQ_VIBHA	P54302 vibrio harv
22	131.5	19.4	441	1 ZRAR_SALTI	Q82333 salmonella
23	131.5	19.4	441	1 ZRAR_SALTY	P58582 salmonella
24	130	19.2	1238	1 EVGS_BORPA	P40330 bordetella
25	128.5	19.0	849	1 LUXN_VIBHA	P54301 vibrio harv
26	127.5	18.8	522	1 MCS4_SCHPO	P87323 schizosacch
27	127	18.7	1238	1 EVGS_BORBR	P26762 bordetella
28	126.5	18.7	227	1 CPXR_HAEIN	P44895 haemophilus
29	125.5	18.5	443	1 ZRAR_KLEOX	Q9apd9 klebsiella
30	124	18.3	144	1 CHBY_TREPA	P6126 treponema p
31	124	18.3	539	1 PRRI_SCHPO	O14283 schizosacch
32	122.5	18.1	441	1 ZRAR_ECOLI	P14375 escherichia
33	122	18.0	248	1 ALGR_PSEAE	P26275 pseudomonas

ALIGNMENTS

RESULT 1

ID	EVGS_ECOLI	STANDARD;	PRT;	1197 AA.
AC	P30855; P77644; Q9RF36; Q9RF37;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Sensor protein evgs precursor (EC 2.7.3.-).			
CN	EVGS OR B2370.			
OS	Escherichia coli			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
EX	MEDLINE=94171083; PubMed=8125343;			
RA	Utsumi R., Katayama S., Taniguchi M., Horie T., Ikeda M., Igaki S., Nakagawa H., Miwa A., Noda M.;			
RA	"Newly identified genes involved in the signal transduction of Escherichia coli K-12.";			
RT	Escherichia coli K-12.";			
RL	Gene 140:73-77(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
EX	MEDLINE=93173621; PubMed=1289796;			
RA	Utsumi R., Katayama S., Ikeda M., Igaki S., Nakagawa H., Miwa A., Taniguchi M., Noda M.;			
RA	"Cloning and sequence analysis of the evgs genes involved in signal transduction of Escherichia coli K-12.";			
RT	Nucleic Acids Symp. Ser. 27:149-150(1992).			
RL				
RN	[3]			
RP	SEQUENCE FROM N.A. (VARIANTS EVGS1 AND EVGS4).			
RC	STRAIN=K12;			
EX	MEDLINE=20378313; PubMed=10923791;			
RA	Kato A., Ohnishi H., Yamamoto K., Furuta E., Tanabe H., Utsumi R.;			
RA	"Transcription of emrKY is regulated by the EvgA-Evgs two-component system in Escherichia coli K-12.";			
RT	Biosci. Biotechnol. Biochem. 64:1203-1209(2000).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12 / MG1655;			
EX	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;			
RA	"The complete genome sequence of Escherichia coli K-12.";			
RT	Science 277:1453-1474(1997).			
RL				
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
EX	MEDLINE=9749980; PubMed=9205837;			
RA	Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K., Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N., Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,			

Q8x613 escherichia
P10957 escherichia
Q34534 bacillus su
Q45994 caulobacter
P16244 escherichia
Q44006 alcaligenes
P49246 xanthomonas
Q56312 thermotoga
P51358 porphyra pu
P43501 pseudomonas
Q06065 escherichia
O78428 guillardia

RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features."; www.ncbi.nlm.nih.gov/pubmed/91113 (1997).
RL DNA Res. 4:91-113 (1997).
RN [6]
RP CHARACTERIZATION.
RX MEDLINE=98194702; PubMed=9535079;
RA Perraud A.-L., Kimmel B., Weiss V., Gross R.;
RT "Specificity of the BvgAS and BvgAS phosphorylation is mediated by the
RT C-terminal Hpt domains of the sensor proteins."; www.ncbi.nlm.nih.gov/pubmed/1071998.
RL Mol. Microbiol. 27:875-887 (1998).
CC -!- FUNCTION: Member of the two-component regulatory system evgs/evga.
CC Phosphorylates evga via a four-step phosphorylation in response to
CC environmental signals.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- PTM: Activation requires a sequential transfer of a phosphate
CC group from a His in the primary transmitter domain, to an Asp in
CC the receiver domain and to a His in the secondary transmitter
CC domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 HPT domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D14008; BAA03108.1; -;
CC EMBL; AF201840; AAF12563.1; -;
CC EMBL; AF201841; AAF17564.1; -;
CC EMBL; AB000325; AAC75429.1; -;
CC EMBL; D30867; BAA16241.1; -;
CC PIR; G65010; G65010.
CC HSSP; P06143; 1UDR.
CC EcoGene; EG11610; evgs.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_Dr_C.
DR InterPro; IPR003661; His_Kinase.
DR InterPro; IPR005467; His_Kinase.
DR InterPro; IPR002570; Hpt.
DR InterPro; IPR001789; Response_reg.
DR InterPro; IPR001311; SBP_glu_Receptor.
DR InterPro; IPR001638; SBP_bac_3.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HiskA; 1.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00497; SBP_bac_3; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HiskA; 1.
DR SMART; SM00073; HPT; 1.
DR SMART; SM00062; PBP; 2.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS10109; HIS_KIN; 1.
DR PROSITE; PS50894; HPT; 1.
DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW Sensory transduction; Transferase; Kinase; Phosphorylation;
KW Sensory transduction; Transferase; Kinase; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 1197
FT DOMAIN 22 325
FT TRANSMEM 326 346
FT DOMAIN 347 537
FT TRANSMEM 538 558
FT POTENTIAL.

FT DOMAIN 559 1197
FT DOMAIN 718 938
FT DOMAIN 960 1074
FT DOMAIN 1098 1197
FT MOD RES 721 721
FT MOD RES 1009 1009
FT MOD RES 1137 1137
FT VARIANT 577 577
FT VARIANT 701 701
FT CONFLICT 152 152
FT CONFLICT 242 242
FT CONFLICT 275 275
FT CONFLICT 420 420
FT CONFLICT 739 739
FT CONFLICT 758 758
FT CONFLICT 761 761
FT CONFLICT 877 877
FT CONFLICT 1045 1045
FT CONFLICT 1074 1074
SQ SEQUENCE 1197 AA; 134742 MW; EBE1DE0F797B1278 CRC64;
Query Match 26.3%; Score 178.5; DB 1; Length 1197;
Best Local Similarity 34.4%; Pred. No. 2.9e-07;
Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;
QY 12 IKKLNVLIVDDPLNLIHEKIKKAGISQANNGEAAVHRDGGSSFDLILDKEM 71
DB 955 LPEKSLIADHDEHTNRLKQLNLLGYDVBATDGVQA--LHKVSMQHYDLITDVNM 1012
QY 72 PERGVSTTKKREMEVKSMIVGVTSIADNEERRAFMEAGLNHCLAKPLTKKIPLIN 131
DB 1013 FNMGGFELTKLRQNSLPIWGLTANA-QANREKGLSCGMNLCFLFKELTDLVLKTHLS 1071
QY 132 QL 133
DB 1072 QL 1073
RESULT 2
EVGS_ECO57 STANDARD; PRT; 1197 AA.
AC P58402;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sensor protein evgs precursor (EC 2.7.3.-).
GN EVGS OR Z3632 OR ECS3249.
OS *Escherichia coli* O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Escherichia*.
OX NCBI_TaxID=83334;
RN [1]_TaxID=83334;
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Kida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba H., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
RT O157:H7 and genomic comparison with a laboratory strain K-12.";

DL DNA Res. 8:11-22(2001).

CC -!- FUNCTION: Member of the two-component regulatory system evgs/evgA.

CC Phosphorylates evgA via a four-step phosphorylation in response to

CC environmental signals (By similarity).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (Probable).

CC -!- PTM: Activation requires a sequential transfer of a phosphate

CC group from a His in the primary transmitter domain, to an Asp in

CC the receiver domain and to a His in the secondary transmitter

CC domain (By similarity).

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

CC -!- SIMILARITY: Contains 1 HPT domain.

CC -!- SIMILARITY: Contains 1 response regulatory domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE005468; AAG57495.1; -;

DR EMBL; AF002561; BAB36672.1; -;

DR PIR; A91035; A91035.

DR PIR; C85879; C85879.

DR InterPro; IPR003594; Atgbind ATPase.

DR InterPro; IPR004358; Bact_sens_pr_C.

DR InterPro; IPR003461; His_KinA.

DR InterPro; IPR005467; His_Kinase.

DR InterPro; IPR002570; Hpt.

DR InterPro; IPR001789; Response_reg.

DR InterPro; IPR001311; SBP_glu_receptor.

DR InterPro; IPR001638; SBP_bac_3.

DR Pfam; PF02518; HtsK_A; 1.

DR Pfam; PF00512; HtsK_A; 1.

DR Pfam; PF00072; response_reg; 1.

DR Pfam; PF00497; SBP_bac_3; 1.

DR PRINTS; PR00344; BCTRLSENSOR.

DR ProDom; PD000039; Response_reg; 1.

DR SMART; SM00387; HATPase_C; 1.

DR SMART; SM00388; HtsK_A; 1.

DR SMART; SM00073; HPT; 1.

DR SMART; SM00062; PBPb; 2.

DR SMART; SM00448; REC; 1.

DR PROSITE; PS0109; HIS_KIN; 1.

DR PROSITE; PS0894; HPT; 1.

DR PROSITE; PS0110; RESPONSE_REGULATORY; 1.

DR Sensory transduction; Transferrase; Kinase; Phosphorylation;

DR Transmembrane; Inner membrane; Signal; Complete proteome.

FT SIGNAL 1 21

FT CHAIN 22 1197

FT DOMAIN 22 325

FT TRANSMEM 326 346

FT DOMAIN 347 537

FT PERIPLASMIC (POTENTIAL).

FT TRANSMEM 538 558

FT DOMAIN 559 1197

FT CYTOPLASMIC (POTENTIAL).

FT HISTIDINE KINASE

FT RESPONSE REGULATORY.

FT DOMAIN 960 1074

FT HPT.

FT MOD_RES 721 721

FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

FT MOD_RES 1009 1009

FT PHOSPHORYLATION (BY SIMILARITY).

FT MOD_RES 1137 1137

FT PHOSPHORYLATION (BY SIMILARITY).

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.

CC -!- SIMILARITY: Contains 1 response regulatory domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

Query Match 26.0%; Score 176.5; DB 1; Length 1197;

Best Local Similarity 34.4%; Pred. No. 4.3e-07;

Matches 42; Conservative 25; Mismatches 52; Indels 3; Gaps 2;

12 IKKLNVLIVDDPDLNLIIHKIKGIGISOTANNGEAVIHRDGGSGFDLILMDKEM 71

955 LPEKLSILIADHPNRLILKRLNQLNGLDYDEATDGVQA--LHKVSMQHYDLITDVM 1012

QY 72 PERDGVSTTKKLEMEVKSMIVGVTSIADNBEERAFMEAGNHGCLAKPLTKDKIIPLIN 131

DB 1013 PNVDGELTKRQEQNSLPINGLTANA-QANERKGLNCGWNLCLFPLTUDVUKTHLS 1071

QY 132 QL 133

DB 1072 QL 1073

RESULT 3

RCSC_SALTI STANDARD; PRT; 948 AA.

AC Q56128;

DT 01-NOV-1997 (Rel. 35, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Sensor protein rcsC (SC 2.7.3.-) (Capsular synthesis regulator

DE component C).

GN RCSC OR STY2496 OR T0594.

OS Salmonella typhi.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=601;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CT18;

RX MEDLINE=21534947; PubMed=11677608;

RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,

RA Baker S., Basham P., Brooks K., Chillingworth T., Connor P.,

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,

RA Feltwell T., Hamlin N., Haque A., Hien T.F., Holroyd S., Jagels K.,

RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,

RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,

RA Whitehead S., Barrell B.G.;

RT "Complete genome sequence of a multiple drug resistant Salmonella

RT enterica serovar Typhi CT18.";

RL Nature 413:848-852(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Ty2 / ATCC 700931;

RX MEDLINE=22531367; PubMed=12644504;

RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,

RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;

RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2

RT and CT18.";

RL J. Bacteriol. 185:2330-2337(2003).

RN [3]

RP SEQUENCE OF 846-948 FROM N.A.

RC STRAIN=Ty2 / ATCC 700931;

RX MEDLINE=96198173; PubMed=8626298;

RA Virlogeux I., Waxin H., Ecobichon C., Lee J.O., Popoff M.Y.;

RT "Characterization of the rcsA and rcsB genes from Salmonella typhi:

RT rcsB through tvia is involved in regulation of Vi antigen

RT synthesis.";

RL J. Bacteriol. 178:1691-1698(1996).

CC -!- FUNCTION: Member of the two-component regulatory system rcsC/rcsB

CC involved in the regulation of the expression of genes involved in

CC colanic acid capsule synthesis. RcsC probably functions as a

CC membrane-associated protein kinase that phosphorylates rcsB in

CC response to environmental signals.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (Probable).

CC -!- PTM: Activation probably requires a transfer of a phosphate group

CC between a His in the transmitter domain and an Asp of the receiver

CC domain (By similarity).

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.

CC -!- SIMILARITY: Contains 1 response regulatory domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

ENBL; AL627274; CAD07502.1; -
 ENBL; A8016836; AAO68299.1; -
 ENBL; X87830; CAA61095.1; -
 HSSP; P06143; IUDR.
 InterPro; IPR003594; ATPbind ATPase.
 InterPro; IPR004358; Bact_sens_pr_C.
 InterPro; IPR003681; His_Kinase.
 InterPro; IPR005467; His_Kinase.
 InterPro; IPR000014; PAS domain.
 InterPro; IPR001789; Response_reg.
 Pfam; PF02518; HATPase_c; 1.
 Pfam; PF00512; Hiska; 1.
 Pfam; PF00072; response_reg; 1.
 PRINTS; PR00344; BCTRLSENSOR.
 ProDom; PD000039; Response_reg; 1.
 SMART; SM00387; HATPase_c; 1.
 SMART; SM00388; Hiska; 1.
 SMART; SM00448; REC; 1.
 PROSITE; PS50109; HIS_KIN; 1.
 PROSITE; PS50112; PAS; FALSE NEG.
 PROSITE; PS50110; RESPONSE REGULATORY; 1.
 Sensory transduction; Transferrase; Kinase; Bacterial capsule;
 Inner membrane; Transmembrane; Phosphorylation; Complete proteome.
 DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 21 41 POTENTIAL.
 DOMAIN 42 313 PERIPLASMIC (POTENTIAL).
 TRANSMEM 314 334 POTENTIAL.
 DOMAIN 335 948 CYTOPLASMIC (POTENTIAL).
 DOMAIN 357 425 PAS.
 DOMAIN 476 692 HISTIDINE KINASE.
 DOMAIN 826 940 RESPONSE REGULATORY.
 MOD_RES 479 479 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 MOD_RES 875 875 PHOSPHORYLATION (BY SIMILARITY).
 SEQUENCE 948 AA; 106237 MW; AE3A21701265A865 CRC64;

Query Match 25.6%; Score 173.5; DB 1; Length 948;
 Best Local Similarity 33.9%; Pred. No. 5.9e-07;
 Matches 37; Conservative 33; Mismatches 36; Indels 3; Gaps 2;
 18 VLIVDDPLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSPDLIMDKEMPERGV 77
 827 ILVVDHPINRLADQLGSLGYCKTANGVDALNVL--SKNAIDVLSDVNMENMDGY 884
 78 STTKLEMEVKSMIVGVTSIADNEERAFMEAGLNHGLAKPLTKDKI 126
 885 RLTORIQRLGLTLPVVGVTANALAEKQRC-LESGMDSCLSKPVTLDAL 932

RESULT 4
 ID RCSC_SALTY STANDARD; PRT; 948 AA.
 AC P58662;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Sensor protein rcsC (EC 2.7.3.-) (Capsular synthesis regulator component C).
 DE component C.
 GN RCSC OR STM2271.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante J., Du F., Hou S., Layman D.,

Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
 "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2.";
 Nature 413:852-856(2001).
 -!- FUNCTION: Member of the two-component regulatory system rcsC/rcsB involved in the regulation of the expression of genes involved in colanic acid capsule synthesis. RcsC probably functions as a membrane-associated protein kinase that phosphorylates rcsB in response to environmental signals.
 -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane. (Probable).
 -!- PTM: Activation probably requires a transfer of a phosphate group between a His in the transmitter domain and an Asp of the receiver domain (By similarity).
 -!- SIMILARITY: Contains 1 histidine kinase domain.
 -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
 -!- SIMILARITY: Contains 1 response regulatory domain.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB008801; AAL21172.1; -
 StyGene; SC77777; rcsC
 InterPro; IPR003594; ATPbind ATPase.
 InterPro; IPR004358; Bact_sens_pr_C.
 InterPro; IPR003681; His_Kinase.
 InterPro; IPR005467; His_Kinase.
 InterPro; IPR000014; PAS domain.
 InterPro; IPR001789; Response_reg.
 Pfam; PF02518; HATPase_c; 1.
 Pfam; PF00512; Hiska; 1.
 Pfam; PF00072; response_reg; 1.
 PRINTS; PR00344; BCTRLSENSOR.
 ProDom; PD000039; Response_reg; 1.
 SMART; SM00387; HATPase_c; 1.
 SMART; SM00388; Hiska; 1.
 SMART; SM00448; REC; 1.
 PROSITE; PS50109; HIS_KIN; 1.
 PROSITE; PS50112; PAS; FALSE NEG.
 PROSITE; PS50110; RESPONSE REGULATORY; 1.
 Sensory transduction; Transferrase; Kinase; Bacterial capsule;
 Inner membrane; Transmembrane; Phosphorylation; Complete proteome.
 DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 21 41 POTENTIAL.
 DOMAIN 42 313 PERIPLASMIC (POTENTIAL).
 TRANSMEM 314 334 POTENTIAL.
 DOMAIN 335 948 CYTOPLASMIC (POTENTIAL).
 DOMAIN 357 425 PAS.
 DOMAIN 476 692 HISTIDINE KINASE.
 DOMAIN 826 940 RESPONSE REGULATORY.
 MOD_RES 479 479 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 MOD_RES 875 875 PHOSPHORYLATION (BY SIMILARITY).
 SEQUENCE 948 AA; 106279 MW; BAAD8DA57D5868B CRC64;

Query Match 25.4%; Score 172.5; DB 1; Length 948;
 Best Local Similarity 34.6%; Pred. No. 7.1e-07;
 Matches 37; Conservative 32; Mismatches 35; Indels 3; Gaps 2;

18 VLIVDDPLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSPDLIMDKEMPERGV 77
 827 ILVVDHPINRLADQLGSLGYCKTANGVDALNVL--SKNAIDVLSDVNMENMDGY 884
 78 STTKLEMEVKSMIVGVTSIADNEERAFMEAGLNHGLAKPLTKD 124
 885 RLTORIQRLGLTLPVVGVTANALAEKQRC-LESGMDSCLSKPVTLD 930

```

RESULT 5
CC ECOLI STANDARD; PRT; 949 AA.
CC P14376; P76457; P97170; P97202; Q47586;
CC 01-JAN-1990 (Rel. 13, Created)
CC 28-FEB-2003 (Rel. 41, Last sequence update)
CC 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sensor protein rcsC (EC 2.7.3.-) (Capsular synthesis regulator
DE component C).
DE NCBI_TaxID=562;
XS Escherichia coli.
XC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
XC Enterobacteriaceae; Escherichia.
XX NCBI_TaxID=562;
XN [1]
XP SEQUENCE FROM N.A.
XC STRAIN=K12;
XX MEDLINE=90130299; PubMed=2404948;
XA Stout V., Gottesman S.
XT "RcsB and RcsC: a two-component regulator of capsule synthesis in
XT Escherichia coli.";
XN J. Bacteriol. 172:659-669 (1990).
XN [2]
XP SEQUENCE FROM N.A.
XC STRAIN=O9:K30:H12;
XX MEDLINE=93374832; PubMed=8366025;
XA Jayaratne P., Keenleyside W.J., MacLachlan P.R., Dodgson C.,
XA Whitfield C.
XT "Characterization of rcsB and rcsC from Escherichia coli O9:K30:H12
XT and examination of the role of the rcs regulatory system in
XT expression of group I capsular polysaccharides.";
XN J. Bacteriol. 175:5384-5394 (1993).
XN [3]
XP SEQUENCE FROM N.A.
XC STRAIN=K12 / MG1655;
XA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
XA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
XA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
XA Mau B., Shao Y.
XT "The complete genome sequence of Escherichia coli K-12.";
XN Science 277:1453-1474 (1997).
XN [4]
XP SEQUENCE FROM N.A.
XC STRAIN=K12;
XX MEDLINE=97251358; PubMed=3097040;
XA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
XA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
XA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
XA Nishimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y.,
XA Sivasubram S., Tagami H., Takeda J., Takemoto K., Wada C.,
XA Yamamoto Y., Horiuchi T.
XT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
XT corresponding to the 40.1-50.0 min region on the linkage map.";
XN DNA Res. 3:379-392 (1996).
CC -!- FUNCTION: Member of the two-component regulatory system rcsC/rcsB
CC involved in the regulation of the expression of genes involved in
CC colanic acid capsule synthesis. RcsC probably functions as a
CC membrane-associated protein kinase that phosphorylates rcsB in
CC response to environmental signals.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- PTM: Activation probably requires a transfer of a phosphate group
CC between a His in the transmitter domain and an Asp of the receiver
CC domain (By similarity).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -!- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 808.
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M28242; AAA4503.1; ALT_INIT.
CC EMBL; L11272; AAA4505.1; ALT_INIT.
CC EMBL; AE003111; AAC75278.1; ALT_INIT.
CC EMBL; D90850; BAA16001.1; ALT_FRAME.
CC EMBL; D90850; BAA16006.1; ALT_FRAME.
CC EMBL; D90851; BAA16009.1; ALT_FRAME.
CC EMBL; D90851; BAA16014.1; ALT_FRAME.
CC HSP; P06143; IUDR.
CC EcGene; EG10822; rcsC.
CC InterPro; IPR003594; ATPbind ATPase.
CC InterPro; IPR004358; Bact_sens_pr_C.
CC InterPro; IPR003661; His_KinA.
CC InterPro; IPR005467; His_kinase.
CC InterPro; IPR000014; PAS_domain.
CC InterPro; IPR001789; Response_reg.
CC Pfam; PF02518; HATPase_c; 1.
CC Pfam; PF00512; HisKA; 1.
CC Pfam; PF00072; response_reg; 1.
CC PRINTS; PR00344; BCTRLSENSOR.
CC ProDom; PD000039; Response_reg; 1.
CC SMART; SM00387; HATPase_c; 1.
CC SMART; SM00388; HisKA; 1.
CC SMART; SM00091; PAS; 1.
CC SMART; SM00448; REC; 1.
CC PROSITE; PSS0109; HIS_KIN; 1.
CC PROSITE; PSS0112; PAS; FALSE NEG.
CC PROSITE; PSS0110; RESPONSE_REGULATORY; 1.
CC Sensory transduction; Transferase; Kinase; Bacterial capsule;
CC Inner membrane; Transmembrane; Phosphorylation; Complete proteome.
CC -----
CC DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 20 41 POTENTIAL.
CC DOMAIN 42 313 PERIPLASMIC (POTENTIAL).
CC TRANSMEM 314 335 POTENTIAL.
CC DOMAIN 336 949 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 357 425 PAS.
CC DOMAIN 476 692 HISTIDINE KINASE.
CC DOMAIN 826 940 RESPONSE REGULATION.
CC MOD RES 479 479 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC MOD RES 875 875 PHOSPHORYLATION (BY SIMILARITY).
CC VARIANT 315 315 L -> V (IN STRAIN O9:K30:H12).
CC VARIANT 776 776 K -> V (IN STRAIN O9:K30:H12).
CC VARIANT 890 890 I -> T (IN STRAIN O9:K30:H12).
CC VARIANT 938 938 L -> V (IN STRAIN O9:K30:H12).
CC VARIANT 948 948 D -> E (IN STRAIN O9:K30:H12).
CC CONFLICT 129 130 MR -> IG (IN REF. 1 AND 2).
CC CONFLICT 935 935 T -> S (IN REF. 1).
CC SEQUENCE 949 AA; 106506 MW; E37E9D70EC944A78 CRC64;
Query Match 25.3%; Score 171.5; DB 1; Length 949;
Best Local Similarity 33.9%; Pred. No. 8.6e-07;
Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;
QY 18 VLIVDDPPLNLIHEKIIKAIGIGISOTANGGEAVIIHRDGGSPFLIMDKEMPERDGV 77
Db 827 ILVVDHPINRLLADQLGSLGYQCKTANDGVDALENVL--SKNHIDIVLSDVNMFMMDGY 884
QY 78 STYTKLREMEVKSMIVGVTSIADNREERAEAFEGNHCIAKPLTKDKI 126
Db 885 RLQIRQLGLTLPVIGVTANALAEKQRC-LESGNDSCLSKRPVLDDVI 932
RESULT 6
ID PRZE MYXXA STANDARD; PRT; 777 AA.
AC P18769;
DT 01-NOV-1990 (Rel. 16, Created)

```

DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gliding motility regulatory protein (EC 2.7.3.-).
 GN PRZE
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cystobacterineae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90332690; PubMed=2165608;
 RA McCleary W.R., Zusman D.R.;
 RT "FrzE of Myxococcus xanthus is homologous to both CheA and CheY of
 RT Salmonella typhimurium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5898-5902(1990).
 RN [2]
 RP PHOSPHORYLATION OF HIS-49.
 RX MEDLINE=91072208; PubMed=2123853;
 RA McCleary W.R., Zusman D.R.;
 RT "Purification and characterization of the Myxococcus xanthus FrzE
 RT protein shows that it has autophosphorylation activity.";
 RL J. Bacteriol. 172:6651-6668(1990).
 CC -!- FUNCTION: FRZE IS INVOLVED IN A SENSORY TRANSDUCTION PATHWAY THAT
 CC CONTROLS THE FREQUENCY AT WHICH CELLS REVERSE THEIR GLIDING
 CC DIRECTION. FRZE SEEMS TO BE CAPABLE OF AUTOPHOSPHORYLATING ITSELF
 CC ON AN HISTIDINE RESIDUE AND THEN TO TRANSFER THAT GROUP TO AN
 CC ASPARTATE RESIDUE IN THE C-TERMINAL PART OF THE PROTEIN.
 CC -!- SIMILARITY: Contains 1 cheW-like domain.
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.
 CC -!- SIMILARITY: Contains 1 HPT domain.
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M35192; AAA25396.1; -.
 DR PIR; A35966; A35966.
 DR HSP; Q36310; LB30.
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR004358; Bact_sens_pr_C.
 DR InterPro; IPR002545; CheW.
 DR InterPro; IPR005467; His_kinase.
 DR InterPro; IPR002570; Hpt.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF01584; CheW; 1.1.
 DR Pfam; PF02518; HATPase_c; 1.
 DR Pfam; PF01627; Hpt; 1.
 DR PRINTS; PR00072; response_reg; 1.
 DR PRINTS; PR00344; BCTRISensor.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00260; CheW; 1.
 DR SMART; SM00387; HATPase_c; 1.
 DR SMART; SM00073; HPT; 1.
 DR SMART; SM00448; REC; 1.
 DR PROSITE; PS50851; CHEW; 1.
 DR PROSITE; PS50109; HIS_KIN; 1.
 DR PROSITE; PS50894; HPT; 1.
 DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
 KW Sensory transduction; Transferase; Kinase; Phosphorylation.
 FT DOMAIN 1 108
 FT HPT
 FT DOMAIN 270 509
 FT HISTIDINE KINASE.
 FT DOMAIN 511 675
 FT CHEW-LIKE.
 FT DOMAIN 660 746
 FT RESPONSE REGULATORY.
 FT MOD_RES 49 49
 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT DOMAIN 130 197
 FT ALA/PRO-RICH (POSSIBLE HINGE REGION).
 SQ SEQUENCE 777 AA; 83189 MW; 9912BD40991C69E5 CRC64;
 Query Match 23.1%; Score 156.5; DB 1; Length 777;
 Best Local Similarity 35.2%; Pred. No. 1.2e-05;
 Matches 38; Conservative 27; Mismatches 38; Indels 5; Gaps 3;
 QY 14 KKLNVLLVDDPLNLIHETIIKAIGISOTANGGEAVIHRDGGSPDLIMDKEMPE 73
 DB 657 KRLRVLLVDDSPARATEGALVKGHSVEEQDGEAYV--KVQNTYDILLITDVQMPK 714
 QY 74 RDGVSTTKKUREMEVKSMI--VGVTSLADNEERRAFMEAGLNHCLAK 119
 DB 715 LDGFSLARLRKSTPAVARIPVILISLASPEDKRRG-LDAGADAYLVK 761
 RESULT 7
 SKN7_YEAST STANDARD; PRT; 622 AA.
 ID SKN7_YEAST
 AC P38889; P39747;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative transcription factor SKN7 (POS9 protein).
 GN SKN7 OR POS9 OR BRV1 OR YHR206W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94042854; PubMed=8226633;
 RA Brown J.L., North S., Bussey H.;
 RT "SKN7, a yeast multicopy suppressor of a mutation affecting cell wall
 RT beta-glucan assembly, encodes a product with domains homologous to
 RT prokaryotic two-component regulators and to heat shock transcription
 RT factors.";
 RL J. Bacteriol. 175:6908-6915(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Krens B., Charizanis C., Entian K.-D.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jiler M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII.";
 RL Science 265:2077-2082(1994).
 RN [4]
 RP FUNCTION, AND MUTAGENESIS.
 RX MEDLINE=95045411; PubMed=7957083;
 RA Brown J.L., Bussey H., Stewart R.C.;
 RT "Yeast Skn7p functions in a eukaryotic two-component regulatory
 RT pathway.";
 RL EMBO J. 13:5186-5194(1994).
 CC -!- FUNCTION: INVOLVED IN OXIDATIVE STRESS. TRANSCRIPTION FACTOR THAT
 CC MAY FUNCTION IN A TWO-COMPONENT SIGNAL TRANSDUCTION PATHWAY THAT
 CC ACTS IN PARALLEL WITH THE PKC1 CASCADE TO REGULATE GROWTH AT THE
 CC CELL SURFACE.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC EMBL: U00485; AAC48911.1; -
DR EMBL: X83031; CA858143.1; -
DR EMBL: U00029; A869734.1; -
DR PIR: A49344; A49344.
DR HSSP: F22121; 2HTS.
DR TRNSPAC: T03481; -.
DR SGD: S0001249; SKN7.
DR GO: GO:0005634; C:nucleus; IDA.
DR GO: GO:0003709; F:transcription factor activity; IDA.
DR GO: GO:0000156; F:two-component response regulator activity; IDA.
DR GO: GO:0006350; P:transcription; IDA.
DR InterPro: IPR002342; HSF DNA BIND.
DR InterPro: IPR002341; HSF ETS.
DR Pfam: PF00447; HSF DNA-bind; 1.
DR Pfam: PF00072; response reg; 1.
DR PRINTS: PR00056; HSFDOMAIN.
DR ProDom: PD001788; HSF DNA-bind; 1.
DR SMART: SM00415; HSF; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS00434; HSF DOMAIN; 1.
DR PROSITE: PS0110; RESPONSE REGULATORY; 1.
DR Transcription regulation; Sensory transduction; Nuclear protein;
KW DNA-binding; Phosphorylation. BY SIMILARITY.
FT DOMAIN 86 190
FT MOD RES 378 492 RESPONSE REGULATORY.
FT MOD RES 427 427 PHOSPHORYLATION (PROBABLE).
FT MUTAGEN 427 427 D->N: DIMINISHED ACTIVITY.
FT MUTAGEN 427 427 D->E: AUGMENTED ACTIVITY.
SQ SEQUENCE 622 AA; 69202 MW; 4C732FD6E326742 CRC64;

Query Match 22.1%; Score 150; DB 1; Length 622;
Best Local Similarity 28.1%; Pred. No. 3.3e-05;
Matches 38; Conservative 32; Mismatches 45; Indels 20; Gaps 4;

QY 12 IKKLVLLVDDPLNLIHKKIKAIGISOTANNGEAEVII---HRDGGSSFDLILMD 68
DB 373 LRKGFHVLVEDDAVSIQCKFLKYGCTVQVSDGLSAISTLEKRV-----YDLVMD 427

QY 69 KEMPERDGVSTTKLREMEVKSMVGVTSVLADNFEERRAFMEAGLNHCLAKPTTKOKI-- 126
DB 428 IVPNPLDGTATSIIVRSFDNETPIIATGIMN-QDLITVQLHGMDILAKPTTRDDLHS 486

QY 127 -----IPLNQ 132
DB 487 ILIRYLKDRIPLCEQ 501

RESULT 8
ARCB_ECO57
ID ARCB_ECO57 STANDARD; PRT; 778 AA.
AC P58363;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aerobic respiration control sensor protein arcb (BC 2.7.3.-).
GN ARCB OR Z4574 OR ECS4089.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Fostel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;

```

```

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2].
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Sainagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL Nature Res. 8:11-22(2001).
CC -!- FUNCTION: Member of the two-component regulatory system arcb/arcbA.
CC Sensor-regulator protein for anaerobic repression of the arc
CC regulon. Activates arcbA via a four-step phosphorylation. ArcB can
CC also dephosphorylate arcbA by a reverse phosphorylation involving His-
CC 717 and Asp-576 (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- PTM: Activation requires a sequential transfer of a phosphate
CC group from a His in the primary transmitter domain, to an Asp in
CC the receiver domain and to a His in the secondary transmitter
CC domain (By similarity).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC -!- SIMILARITY: Contains 1 HFT domain.
CC -!- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: AE005549; AAC58344.1; -.
CC EMBL: AP002564; BAB37512.1; -.
CC PIR: A91140; A91140.
CC PIR: D85985; D85985.
CC InterPro: IPR003594; ATPbind_ATPase.
CC InterPro: IPR003661; His_kinA.
CC InterPro: IPR005467; His_kinase.
CC InterPro: IPR002570; Hpt_kinase.
CC InterPro: IPR000700; PAS-associ.
CC InterPro: IPR000014; PAS_domain.
CC InterPro: IPR001789; Response_reg.
CC Pfam: PF02518; HATPase_C; 1.
CC Pfam: PF00512; HiskA; 1.
CC Pfam: PF00989; PAS; 1.
CC Pfam: PF00072; response_reg; 1.
CC ProDom: PD000039; Response_reg; 1.
CC SMART: SM00387; HATPase_c; 1.
CC SMART: SM00388; HiskA; 1.
CC SMART: SM00073; HPT; 1.
CC SMART: SM00091; PAS; 1.
CC SMART: SM00448; REC; 1.
CC TIGRFAMs: TIGR00229; sensory_box; 1.
CC PROSITE: PS0109; HIS_KIN; 1.
CC PROSITE: PS0694; HPT; 1.
CC PROSITE: PS0113; PAC; 1.
CC PROSITE: PS0112; PAS; 1.
CC PROSITE: PS0110; RESPONSE REGULATORY; 1.
CC Sensory transduction; Transference; Kinase; Phosphorylation;
KW Transmembrane; Inner membrane; Transcription regulation;
KW Complete proteome.
FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 26 46 POTENTIAL.
FT DOMAIN 47 57 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).

```

```
FT DOMAIN 153 223 PAS.
FT DOMAIN 226 278 PAC.
FT DOMAIN 289 507 HISTIDINE KINASE.
FT DOMAIN 527 643 RESPONSE REGULATORY.
FT DOMAIN 678 771 HPT.
FT MOD_RES 292 292 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 576 576 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 717 717 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 778 AA; 88010 MW; C8AE004B007F9D30 CRC64;

Query Match 22.1%; Score 149.5; DB 1; Length 778;
Best Local Similarity 25.6%; Pred. No. 4.6e-05;
Matches 31; Conservative 40; Mismatches 47; Indels 2;

QY 16 LNVLIYDDPLNLIHKEIKIAGISQTANNGBEAVIHRDGGSSFDLIIMDKEMPERD 75
DB 526 LNVLIYDDPLNLIHKEIKIAGISQTANNGBEAVIHRDGGSSFDLIIMDKEMPERD 583
QY 76 GVSTTKKLEMEVKSMIVGTSIADNE--EERRAFNEAGLNHCLAKPITKDKIIPILNQLM 134
DB 584 GLDISRELTKYRPEDLPPLVALTANVLKDKQEYLNAGWDVLSKPLSVPALTMIKFW 643
QY 135 D 135
DB 644 D 644

RESULT 9
ARCH_ECOLI STANDARD; PRT; 778 AA.
ID ARCH_ECOLI STANDARD; PRT; 778 AA.
AC P22763;
DT 01-AUG-1991 (Rel. 19, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Aerobic respiration control sensor protein arcbB (EC 2.7.3.-).
GN ARCB OR B3210 OR SP3250.
OS Escherichia coli, and
CS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CX NCBI_TaxID=562, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=9035583; PubMed=2201868;
RA Iuchi S., Matsuda Z., Fujiwara T., Lin E.C.C.;
RT "The arcb gene of Escherichia coli encodes a sensor-regulator protein
for anaerobic repression of the arc modulon."
RL Mol. Microbiol. 4:715-727(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
May B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384530;
RA Jin Q., Yuan J., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [4]
RP CHARACTERIZATION.
```

```
RC SPECIES=E.coli; STRAIN=M15;
RX MEDLINE=97431492; PubMed=9286997;
RA Georgellis D., Lynch A.S., Lin E.C.C.;
RT "In vitro phosphorylation study of the arc two-component signal
transduction system of Escherichia coli."
RL J. Bacteriol. 179:5429-5435(1997).
RN [5]
RP CHARACTERIZATION.
RC SPECIES=E.coli; STRAIN=M15;
RX MEDLINE=99047671; PubMed=9830034;
RA Georgellis D., Kwon O., De Wulf P., Lin E.C.C.;
RT "Signal decay through a reverse phosphorelay in the arc two-component
signal transduction system."
RL J. Biol. Chem. 273:32864-32869(1998).
RN [6]
RP MUTAGENESIS OF HIS-292; ASP-576 AND HIS-717.
RC SPECIES=E.coli; STRAIN=K12 / MC4100;
RX MEDLINE=20309722; PubMed=10851007;
RA Kwon O., Georgellis D., Lin E.C.C.;
RT "Phosphorelay as the sole physiological route of signal transmission
by the arc two-component system of Escherichia coli."
RL J. Bacteriol. 182:3858-3862(2000).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS) OF 660-778.
RC SPECIES=E.coli;
RX MEDLINE=97207018; PubMed=9054511;
RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
RT "Insights into multistep phosphorelay from the crystal structure of
the C-terminal HPT domain of ArcB."
RL Cell 88:717-723(1997).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 659-776 IN COMPLEX WITH CHEY.
RC SPECIES=E.coli;
RX MEDLINE=98437504; PubMed=9761838;
RA Kato M., Mizuno T., Hakoshima T.;
RT "Crystallization of a complex between a novel C-terminal transmitter,
HPT domain, of the anaerobic sensor kinase ArcB and the chemotaxis
response regulator CheY."
RL Acta Crystallogr. D 54:140-142(1998).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 659-776.
RC SPECIES=E.coli;
RX MEDLINE=20003135; PubMed=10531481;
RA Kato M., Mizuno T., Shimizu T., Hakoshima T.;
RT "Refined structure of the histidine-containing-phosphotransfer (HPT)
domain of the anaerobic sensor kinase ArcB from Escherichia coli at
1.57-A resolution."
RL Acta Crystallogr. D 55:1842-1849(1999).
RN [10]
RP FUNCTION: Member of the two-component regulatory system arcbB/arcbA.
Sensor-regulator protein for anaerobic repression of the arc
modulon. Activates arcbA via a four-step phosphorelay. ArcB can
also dephosphorylate arcbA by a reverse phosphorelay involving His-
717 and Asp-576.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(PProbable).
CC -1- PTM: Activation requires a sequential transfer of a phosphate
group from a His in the primary transmitter domain, to an Asp in
the receiver domain and to a His in the secondary transmitter
domain.
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
CC -1- SIMILARITY: Contains 1 HPT domain.
CC -1- SIMILARITY: Contains 1 PAS (PER-ARNT-SIM) dimerization domain.
CC -1- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -1- SIMILARITY: Contains 1 response regulatory domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```


DR EMBL; X53315; CAA37397.1; --
 DR EMBL; U18997; AAS8012.1; --
 DR EMBL; AE000400; AAC76242.1; --
 DR EMBL; AE015336; AAN44715.1; --
 DR FTR; D65112; RGCAR.
 DR PDB; 1A0B; 18-MAR-98.
 DR PDB; 2A0B; 17-JUN-98.
 DR PDB; 1BDJ; 11-MAY-99.
 DR PDB; 1FRO; 31-DEC-02.
 DR CGene; EG10062; arcB.
 DR InterPro; IPR003594; AtPbind_ATPase.
 DR InterPro; IPR004358; Bact_sens_pr_C.
 DR InterPro; IPR003661; His_KinA.
 DR InterPro; IPR005467; His_Kinase.
 DR InterPro; IPR002570; Hpt.
 DR InterPro; IPR000700; PAS-associ C.
 DR InterPro; IPR000014; PAS domain.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF02518; HATPase_C; 1.
 DR Pfam; PF00512; HiskA; 1.
 DR Pfam; PF00989; PAS; 1.
 DR Pfam; PF00072; response_reg; 1.
 DR PRINTS; PR00344; BCTRLSENSOR.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00388; HiskA; 1.
 DR SMART; SM00093; HPT; 1.
 DR SMART; SM00091; PAS; 1.
 DR SMART; SM00448; REC; 1.
 DR TIGRFAMs; TIGR00229; sensory_box; 1.
 DR PROSITE; PS50109; HPT_KIN; 1.
 DR PROSITE; PS50894; HPT; 1.
 DR PROSITE; PS50113; PAC; 1.
 DR PROSITE; PS50112; PAS; 1.
 DR PROSITE; PS50110; RESPONSE_REGULATORY; 1.
 KW Sensory transduction; Transferase; Kinase; Phosphorylation;
 KW Transmembrane; Inner membrane; Transcription regulation; 3D-structure;
 KW Complete proteome.
 FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 26 46 POTENTIAL.
 FT DOMAIN 47 57 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 58 78 POTENTIAL.
 FT DOMAIN 79 778 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 153 223 PAS.
 FT DOMAIN 226 278 PAC.
 FT DOMAIN 289 507 HISTIDINE KINASE.
 FT DOMAIN 527 643 RESPONSE REGULATORY.
 FT DOMAIN 678 771 HPT.
 FT MOD_RES 292 292 PHOSPHORYLATION (AUTO-).
 FT MOD_RES 576 576 PHOSPHORYLATION (PROBABLE).
 FT MOD_RES 717 717 PHOSPHORYLATION (PROBABLE).
 FT MUTAGEN 292 292 H->Q: LOSS OF ACTIVITY.
 FT MUTAGEN 576 576 D->A: LOSS OF ACTIVITY.
 FT MUTAGEN 717 717 H->Q: LOSS OF ACTIVITY.
 FT CONFLICT 469 470 MISSING (IN REF. 2).
 SQ SEQUENCE 778 AA; 87982 MW; D661EAGSGF9AD30 CRC64;

Query Match 22.1%; Score 149.5; DB 1; Length 778;

Best Local Similarity 25.6%; Pred. No. 4.6e-05;

Matches 31; Conservative 40; Mismatches 47; Indels 3; Gaps 2;

16 LNVLLVDDPDLNLIHEKIIKXGIGQITANNGEAVIHRDGGSPFDLIMDKENPERD 75

526 LNVLLVDDPDLNLIHEKIIKXGIGQITANNGEAVIHRDGGSPFDLIMDKENPERD 583

76 GYSTTKKLEMEKSMIVGVTSLADNE-EERRAFMEAGLNHCLAKPLTKDIPLINQLM 134

584 GLDISRELTKRYPREDLPLVALTANVLKDKQKQYLNAGMDVLLSKPLSVLPALTAMIKKFW 643

135 D 135

644 D 644

RESULT 10
 ID SLN1_YEAST STANDARD; PRT; 1220 AA.
 AC P39928;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Osmolarity two-component system protein SLN1 (EC 2.7.3.-).
 GN SLN1 OR YPD2 OR YII147C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / YPH1;
 RX MEDLINE=94024010; PubMed=8211183;
 RA Ota I.M., Varshavsky A.;
 RT "A yeast protein similar to bacterial two-component regulators.";
 RL Science 262:566-569(1993).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169870;
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
 RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
 RL Nature 387:84-87(1997).
 [3]
 RP MUTAGENESIS OF HIS-576 AND ASP-1144.
 RX MEDLINE=94239498; PubMed=8183345;
 RA Maeda T., Wurgler-Murphy S.M., Saito H.;
 RT "A two-component system that regulates an osmosensing MAP kinase
 RT cascade in yeast.";
 RL Nature 369:242-245(1994).
 CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM
 CC SLN1/SSK1 ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE
 CC EXTRACELLULAR ENVIRONMENT. THIS SYSTEM CONTROLS THE SSK2/SSK22->
 CC PBS2-HOG1 PATHWAY. INACTIVE SLN1 ALLOWS THE UNPHOSPHORYLATED SSK1
 CC PROTEIN TO ACTIVATE SSK2 AND SSK22, TWO MAPKKS THAT FURTHER
 CC STIMULATE THE PBS2-HOG1 MAPK CASCADE. IN LOW OSMOLARITY MEDIA,
 CC THE ACTIVATED SLN1 HISTIDINE KINASE REPRESSSES THE ACTIVATION OF THE
 CC PBS2-HOG1 KINASE CASCADE THROUGH PHOSPHORYLATION OF SSK1.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- PTM: ACTIVATION PROBABLY REQUIRES A TRANSFER OF A PHOSPHATE GROUP
 CC BETWEEN A HIS IN THE TRANSMITTER DOMAIN AND AN ASP OF THE RECEIVER
 CC DOMAIN.
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z38059; CAA86131.1; --
 CC EMBL; U01835; AAC48912.1; --
 CC FTR; S48387; S48387.
 CC HSSP; P06143; 1UDR.
 CC SGD; S0001409; SLN1.
 CC GO; GO:0007234; P:osmosensory signaling pathway via two-compo. .; IDA.
 CC GO; GO:0042542; P:response to hydrogen peroxide; IMP.
 CC InterPro; IPR003594; AtPbind_ATPase.
 CC InterPro; IPR004358; Bact_sens_pr_C.
 CC InterPro; IPR003661; His_KinA.
 CC InterPro; IPR005467; His_Kinase.
 CC InterPro; IPR001789; Response_reg.


```

DR Pfam: PF00518; HATPase_c; 1.
DR Pfam: PF00512; HspA; 1.
DR Pfam: PF00072; response_reg; 1.
DR PRINTS: PR00344; BTRLSSENSOR.
DR ProDom: PD000039; Response_reg; 1.
DR SMART: SM00387; Harpase_c; 1.
DR SMART: SM00388; HspA; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
DR Sensory transduction; Transferase; Kinase; Phosphorylation;
KW Transmembrane.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 46 POTENTIAL.
FT DOMAIN 47 333 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 334 354 POTENTIAL.
FT DOMAIN 355 1220 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 573 928 HISTIDINE KINASE.
FT DOMAIN 1089 1210 RESPONSE REGULATORY.
FT MOD_RES 576 576 PHOSPHORYLATION (AUTO-) (PROBABLE).
FT MOD_RES 1144 1144 PHOSPHORYLATION (PROBABLE).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 224 224 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT MUTAGEN 576 576 H->Q: INACTIVE.
FT MUTAGEN 891 891 G->D: SLOW GROWTH, SLN1-1 MUTANT.
FT MUTAGEN 1144 1144 D->N: INACTIVE.
SQ SEQUENCE 1220 AA; 134434 MW; 45FFE24A8165486B CRC64;

Query Match 21.4%; Score 145; DB 1; Length 1220;
Best Local Similarity 29.0%; Pred. No. 0.00018;
Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;

QY 8 DISKIKKLVIVDDPPLNLIHEKIIKA---IGGIS--QTANNGEA---VTHRDGG 59
Db 1080 DDDKNETSVKILVEDNHVN---QEVKRLNLGIEINIELACDQGFAPKVKELTSKG 1135
QY 60 SSFDLILDKEMPERGVSTTKKL-REMEVKSNIYGVTSGLADNEERAFMEAGLNHCLA 118
Db 1136 ENYMLFMVQMPKVDGLLSTMRIRDLGYTFVLTAFADDSNIKEC-LESGNGFLS 1194
QY 119 KPLTKDKIPLNLQMDA 136
Db 1195 KPIKPKLNTILFECAA 1212

RESULT 11
SSK1_YEAST STANDARD; PRT; 712 AA.
AC Q07084; Q07909;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Osmolarity two-component system protein SSK1.
GN SSK1 OR YLR006C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RX MEDLINE=94239498; PubMed=8181345;
RA Maeda T., Wurgler-Murphy S.M., Saito H.;
RT "A two-component system that regulates an osmosensing MAP kinase
RL cascade in yeast."
RL Nature 369:242-245(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313267; PubMed=9169871;

```

```

RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoerg W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Coffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miesga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Paravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM
CC -!- SLN1/SSK1 ACTIVATED BY CHANGES IN THE OSMOLARITY OF THE
CC EXTRACELLULAR ENVIRONMENT. THIS SYSTEM CONTROLS THE SSK2/SSK22->
CC PBS2->HOG1 PATHWAY. THE UNPHOSPHORYLATED FORM OF SSK1 ACTIVATES
CC SSK2 AND SSK22. TWO MAPKKS THAT FURTHER STIMULATE THE PBS2-HOG1
CC MARK CASCADE. IN LOW-OSMOLARITY MEDIA, IS INHIBITED THROUGH THE
CC PHOSPHORYLATION BY SLN1.
CC -!- SUBUNIT: SEEMS TO INTERACT WITH SSK2 AND SSK22.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: L26523; AAA35100.1; -.
CC EMBL: 273178; CAB97528.1; -.
CC DDB: S64828; S64828.
CC SCD: S0003996; SSK1.
CC GO: GO:0007234; P: osmosensory signaling pathway via two-compo. . . ; IDA.
CC GO: GO:0042543; P: response to hydrogen peroxide; IMP.
CC InterPro: IPR001789; Response reg.
CC Pfam: PF00072; response reg; 1.
CC ProDom: PD000039; Response_reg; 1.
CC SMART: SM00448; REC; 1.
CC PROSITE: PS0110; RESPONSE_REGULATORY; 1.
KW Sensory transduction; Phosphorylation.
FT DOMAIN 505 647 RESPONSE REGULATORY.
FT MOD_RES 554 554 PHOSPHORYLATION (PROBABLE).
FT MUTAGEN 554 554 D->N: ACTIVATES.
FT CONFLICT 181 181 P -> S (IN REF. 1).
SQ SEQUENCE 712 AA; 78529 MW; 33B2DBB4FCF2528A CRC64;

Query Match 21.2%; Score 143.5; DB 1; Length 712;
Best Local Similarity 27.9%; Pred. No. 0.00013;
Matches 39; Conservative 28; Mismatches 42; Indels 31; Gaps 3;

QY 10 EKIKKKLVIVDDPPLNLIHEKIIKAIGGISQTANNGEAIVTHRDGSSFDLILMDK 69
Db 498 EKVPKINVLIVDENVINQALGSLFKHKISYKLAKNGQEAIVNWKGG--LHLIFMDL 555
QY 70 EMPERDGVSTTKGREMEVKS-----MIVGVTSLADN 101
Db 556 QLPVLGSGIEAAKQIRDFKQNGIGIQKSLNNSHENLEKGTSKRFSQAPVIVALTASNSQ 615
QY 102 EEERAFMEAGLNHCLAKPL 121
Db 616 MDRKALL-SGCDNYLTQPV 634

RESULT 12
GACS_PSESY STANDARD; PRT; 907 AA.
ID_GACS_PSESY
AC P48027;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

```

28-FEB-2003 (Rel. 41, Last annotation update)
 Sensor protein gacs (EC 2.7.3.-).
 GACS OR LEVA.
 Pseudomonas syringae (pv. syringae).
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=321;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=92234961; PubMed=1314807;
 Hrabak E.M., Willis D.K.;
 "The lea gene required for pathogenicity of Pseudomonas syringae pv.
 syringae on bean is a member of a family of two-component
 regulators.";
 J. Bacteriol. 174:3011-3020(1992).
 CC -!- FUNCTION: FORMS PART OF A TWO-COMPONENT REGULATORY SYSTEM
 GACS/GACS (LEVA). MAY BE INVOLVED IN LESION FORMATION, SWARMING AND
 IN THE PRODUCTION OF EXTRACELLULAR PROTEASE, SYRINGOMYCIN AND N-
 ACYL-L-HOMOSERINE LACTONE (ACYL-HSL). REQUIRED FOR PATHOGENICITY
 ON BEAN.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (Probable).
 CC -!- PTM: Activation requires a sequential transfer of a phosphate
 group from a His in the primary transmitter domain, to an Asp in
 the receiver domain and to a His in the secondary transmitter
 domain (By similarity).
 CC -!- SIMILARITY: Contains 1 HAMP domain.
 CC -!- SIMILARITY: Contains 1 histidine kinase domain.
 CC -!- SIMILARITY: Contains 1 HPT domain.
 CC -!- SIMILARITY: Contains 1 response regulatory domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M80477; AAA25877.1; -;
 CC HSSP; P06143; LUDR
 CC InterPro; IPR003594; ATPbind_ATPase.
 CC InterPro; IPR004358; Bact_sens_pr_C.
 CC InterPro; IPR003660; HAMP.
 CC InterPro; IPR003661; His_kinA.
 CC InterPro; IPR005457; His_kinase.
 CC InterPro; IPR002570; Hpt_kinase.
 CC InterPro; IPR001789; Response_reg.
 CC Pfam; PF00672; HAMP; 1.
 CC Pfam; PF02518; HATPase_c; 1.
 CC Pfam; PF00512; Hiska; 1.
 CC Pfam; PF01627; Hpt; 1.
 CC Pfam; PF00072; response_reg; 1.
 CC PRINTS; PR00344; BCTRLSENSOR.
 CC ProDom; PD000039; Response_reg; 1.
 CC SMART; SM00304; HAMP; 1.
 CC SMART; SM00387; HATPase_c; 1.
 CC SMART; SM00388; Hiska; 1.
 CC SMART; SM00448; REC; 1.
 CC SMART; SM00073; HPT; 1.
 CC PROSITE; PS00885; HAMP; 1.
 CC PROSITE; PS0109; HIS_KIN; 1.
 CC PROSITE; PS00894; HPT; 1.
 CC PROSITE; PS0110; RESPONSE_REGULATORY; 1.
 CC Sensory transduction; Transferase; Kinase; Phosphorylation;
 CC Transmembrane; Inner membrane.
 CC TRANSMEM 9 25 POTENTIAL.
 CC TRANSMEM 84 101 POTENTIAL.
 CC TRANSMEM 159 178 POTENTIAL.
 CC DOMAIN 182 234 HAMP.
 CC DOMAIN 281 502 HISTIDINE KINASE.
 CC DOMAIN 658 777 RESPONSE REGULATORY.
 CC DOMAIN 814 907 HPT.
 CC -----

FT MOD_RES 284 284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 707 707 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 853 853 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 907 AA; 99195 MW; 559F4663DAF3492C CRC64;
 Query Match 20.9%; Score 142; DB 1; Length 907;
 Best Local Similarity 27.6%; Pred. No. 0.00023;
 Matches 35; Conservative 33; Mismatches 33; Indels 26; Gaps 4;
 QY 18 VLIIVDDPPLNLIHEKIKAIGGISQTANNGBEAVIHRDGG-----SSFDILMD 68
 DB 659 VLVDDPPLNLIHQIIEKIGISQTANNGBEAVIHRDGG-----SSFDILMD 707
 QY 69 KEMPERGVSTTKKLEMEVKS-----MIVGVTSLDNNEERAFMEAGLNHCLAPLTK 123
 DB 708 VQPMGMDGRQATEAIRAEAEARNQSSLPFIVALTAAH-MANEKRSILQSGMDDYLTKEISE 766
 QY 124 DKTIPLI 130
 DB 767 RLQAQVV 773
 RESULT 13
 DCTR BACSU
 ID DCTR BACSU STANDARD; PRT; 226 AA.
 AC P96502; P94503;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable C4-dicarboxylate response regulator dctr.
 GN DCTR.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kasahara Y., Nakai S., Lee S., Sadaie Y., Ogasawara N.;
 RT "A 148 kbp sequence of the region between 35 and 47 degree of the
 RL Bacillus subtilis genome";
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=980404033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maubert C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Fumelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler H., Wedler H., Weitzengger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

XX MEDLINE=91286247; PubMed=1905718;

YA Bischoff D.S., Ordal G.W.;

YB "Sequence and characterization of *Bacillus subtilis* CheB, a homolog
YC of *Escherichia coli* CheY, and its role in a different mechanism of
YD chemotaxis.";

YF J. Biol. Chem. 266:12301-12305(1991).

YH (2)

YI SEQUENCE FROM N.A.

YJ STRAIN=168;

YK MEDLINE=98044033; PubMed=3384377;

YL Kunet F., Ogasawara N., Moser J., Albertini A.M., Alloni G.,
YM Azevedo V., Berrero M.G., Bessieres P., Bolotin A., Borchert S.,
YN Broussier R., Bourrier L., Brans A., Braun M., Brignell S.C., Bron S.,
YO Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
YQ Choi S.X., Codani J.J., Conerton I.P., Cummings N.J., Daniel R.A.,
YR Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
YS Enian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
YT Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
YU Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
YV Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Renaut A.,
YW Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
YX Joris B., Karamata D., Kaashara Y., Klaerr-Blanchard M., Klein C.,
YY Kobayashi Y., Koetter P., Koningsstein G., Krohn S., Kumano M.,

ZA Kurita K., Lepidus A., Lardinois S., Lauber J., Lazarevic V.,

ZB Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,

ZC Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

ZE Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

ZF Parro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,

ZG Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

ZH Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,

ZI Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

ZJ Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,

ZK Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,

ZL Takeuchi M., Tamakoshi A., Tanaka T., Terpatra P., Tognoni A.,

ZM Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,

ZN Viari A., Wambutt R., Wedler E., Wedler H., Weitzensagger T.,

ZO Winters P., Wipat R., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

ZA Yoshida K., Yoshikawa H.F., Zumbach H., Yoshikawa H., Danchin A.,

ZB "The complete genome sequence of the Gram-positive bacterium *Bacillus*
ZC *subtilis*.";

ZD Nature 390:249-256(1997).

ZE (3)

ZF SEQUENCE OF 1-7 FROM N.A.

ZG STRAIN=168 / OI1085;

ZH MEDLINE=93078625; PubMed=1447979;

ZI Bischoff D.S., Ordal G.W.;

ZJ "Identification and characterization of *FlhY*, a novel component of
ZK the *Bacillus subtilis* flagellar switch complex.";

ZL Mol. Microbiol. 6:2715-2723(1992).

ZM (4)

ZN SEQUENCE OF 95-119 FROM N.A.

ZO STRAIN=168 / OI1085;

ZA MEDLINE=92283757; PubMed=1597417;

ZB Bischoff D.S., Weinreich M.D., Ordal G.W.;

ZC "Nucleotide sequences of *Bacillus subtilis* flagellar biosynthetic
ZD genes *flp* and *fliQ* and identification of a novel flagellar gene,
ZE *fliz*.";

ZF (5)

ZG J. Bacteriol. 174:4017-4025(1992).

ZH (5)

ZI SEQUENCE OF 1-12.

ZJ STRAIN=168 / JH642;

ZK MEDLINE=96345629; PubMed=8755892;

ZL Gramann P., Schroeder K., Schmid R., Marahel M.A.;

ZM "Cold shock stress-induced proteins in *Bacillus subtilis*.";

ZN J. Bacteriol. 178:4611-4619(1996).

ZO -!- FUNCTION: HAS A CENTRAL CONTROLLING ROLE IN CHEMOTAXIS. ALTHOUGH
ZI IT IS HOMOLOGOUS TO THE CHEY PROTEIN OF OTHER BACTERIA, IT IS
ZJ THOUGHT TO FUNCTION IN A DIFFERENT MANNER.

ZK -!- SUBCELLULAR LOCATION: Cytoplasmic.

ZL -!- SIMILARITY: Contains 1 response regulatory domain.

ZM (5)

ZN This SWISS-PROT entry is copyright. It is produced through a collaboration
ZO between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M59781; AAA22311.1; -

CC EMBL; Z99112; CAB13506.1; -

CC EMBL; M86738; AAA22450.1; -

CC EMBL; M87005; AAA22451.1; -

CC PIR; A40874; A40874.

CC HSSP; Q56312; ITMY.

CC Subtilist; BG10258; cheY.

CC InterPro; IPR001789; Response reg.

CC Pfam; PF00072; response reg. 1.

CC ProDom; PD000039; Response_reg; 1.

CC SMART; SM00448; REC; 1.

CC FROSITE; FS00110; RESPONSE REGULATORY; 1.

CC Chemotaxis; Sensory transduction; Complete proteome.

CC INIT MET 0

CC DOMAIN 1 118 RESPONSE REGULATORY.

CC MOD RES 53 53 PHOSPHORYLATION (BY SIMILARITY).

CC SEQUENCE 119 AA; 13178 MW; F3BCA0F02CAB7531 CRC64;

CC Query Match 19.8%; Score 134; DB 1; Length 119;

CC Best Local Similarity 26.9%; Pred. No. 0.00012;

CC Matches 32; Conservative 36; Mismatches 47; Indels 4; Gaps 3;

CC 18 VLIVDDDP-LNLIHKKIKAGISQITANNGEAAVHHRDGGSSFDLILMDKEMPERDG 76

CC 4 ILIVDDDAFWRMKILVKNQGEVVAENGAQAVEKYKE--HSPDLVTMDITWPEVDG 61

CC 77 VSTTKLREMEVSMIVGVTSVLADNEERAEAFWEAGLNELAKPLTKKILPLINQMD 135

CC 62 ITALKEIKQIDAQARIIMCSAMGQSQSMVIDA-IQAGAKDFIVKPFQADRVLEINKTLN 119

Search completed: August 14, 2003, 16:47:19

Job time : 24 secs

Db 869 RLQRIQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDVI 916

RESULT 9

E85861

hypothetical protein rscC [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: E85861

R:Perna, N.T.; Plunkett III, G.; Butland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, R.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E85861

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-933 <STO>

A:Cross-references: GB:AE005174; NID:g12516556; PIDN:AA057353.1; GSPDB:GN00145; UWGP:Z34

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

C:Superfamily: rscC protein; response regulator homology

Query Match 25.3%; Score 171.5; DB 2; Length 933;

Best Local Similarity 33.9%; Pred. No. 1.5e-06;

Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;

Qy 18 LVIVDDPDLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLIMDKEMPERDV 77

Db 811 LVVDDHPINRLADQLGSLGYQCTANDGVDAENVL--SKNHIDIVLSNVNPNMDGY 868

Qy 78 STTKKLEMEVKSMIVGVTSIADNEERRAFMEAGLNHCLAKPLTKDKI 126

Db 869 RLQRIQLGLTLPVIGVTANALAEKQRC-LESGMDSCLSKPVTLDVI 916

RESULT 10

AD0149

two component sensor kinase/response regulator protein RscC (EC 2.7.3.-) [imported] - Yersinia pestis

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001

C:Accession: AD0149

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0149

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-957 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90055.1; PID:g15979276; GSPDB:GN00175

C:Genetics:

A:Gene: rscC

C:Superfamily: rscC protein; response regulator homology

C:Keywords: phosphotransferase

Query Match 25.3%; Score 171.5; DB 2; Length 957;

Best Local Similarity 35.1%; Pred. No. 1.5e-06;

Matches 39; Conservative 29; Mismatches 40; Indels 3; Gaps 2;

Qy 16 LVIVDDPDLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLIMDKEMPERD 75

Db 835 LQIVVDDHPINRLADQLTIGVITANDGLDAIVALT--NTVDVILVDVNMVND 892

Qy 76 GVSTTKKLEMEVKSMIVGVTSIADNEERRAFMEAGLNHCLAKPLTKDKI 126

Db 893 GYRLTERLRQLNHNPIIGVTANALAEKQRC-IEAGMNDCLSKPVTLDLT 942

RESULT 11

E87644

sensory box histidine kinase/response regulator [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: E87644

R:Nierman, W.C.; Deiblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo, n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87644

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-642 <STO>

A:Cross-references: GB:AE005673; NID:g13424865; PIDN:AAK25153.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC3191

Query Match 24.0%; Score 162.5; DB 2; Length 642;

Best Local Similarity 33.0%; Pred. No. 5.5e-06;

Matches 36; Conservative 28; Mismatches 40; Indels 5; Gaps 2;

Qy 16 LVIVDDPDLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLIMDKEMPERD 75

Db 514 LVLLVVDNATNRLIATRMLEALGARVTTAEDGAGVAAARQG---FDLLIMDIQMPVMD 570

Qy 76 GYSTTKKLEMEVKSMIVGVTSIADN--EERRAFMEAGLNHCLAKPLT 122

Db 571 GYEATHRIAFNSPAGAPILAMTANAWHQASVLAAGMDGAIKAPLS 619

RESULT 12

AD0432

aerobic respiration control sensor/response regulatory protein (EC 2.7.3.-) [imported]

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001

C:Accession: AD0432

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AD0432

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-778 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC92784.1; PID:g15981477; GSPDB:GN00175

C:Genetics:

A:Gene: arcB

C:Superfamily: aerobic respiration control sensor protein arcB; response regulator homo

C:Keywords: phosphotransferase

Query Match 24.0%; Score 162.5; DB 2; Length 778;

Best Local Similarity 27.0%; Pred. No. 6.8e-06;

Matches 33; Conservative 38; Mismatches 48; Indels 3; Gaps 2;

Qy 16 LVIVDDPDLNLIHEKIIKAIGISQTANNGEAVIIHRDGGSSFDLIMDKEMPERD 75

Db 525 LVLLVVDNATNRLIATRMLEALGARVTTAEDGAGVAAARQG---EEDFVLVDLQIDPMS 582

Qy 76 GYSTTKKLEMEVKSMIVGVTSIADN--EERRAFMEAGLNHCLAKPLTKDKI 134

Db 583 GLDIARQIRAEYKQSLPPLVALTANVLKDKKVEDLAGMDVLSKPLSVPAITAMIKQFW 642

Qy 135 DA 136

Db 643 DS 644

RESULT 13

A69487

Search completed: August 14, 2003, 16:51:17
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:49:07 ; Search time 55 Seconds
(without alignments)
323.935 Million cell updates/sec

Title: US-09-646-679-15
Perfect score: 678
Sequence: 1 MATKSMGDIKXKKLVLI.....LAKPLTKDKIPLINQLMDA 136

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 492763 seqs, 131003257 residues

Total number of hits satisfying chosen parameters: 492763

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	202	29.8	261	15	US-10-101-464A-127
2	171.5	25.3	118	10	US-09-918-508-8
3	150	22.1	622	10	US-09-801-368-328
4	146	21.5	1018	15	US-10-101-464A-909
5	145	21.4	1220	10	US-09-801-368-332
6	143.5	21.2	712	10	US-09-801-368-366
7	137.5	20.3	974	15	US-10-126-120-2
8	137	20.2	125	10	US-09-918-508-7
9	132	19.5	1829	15	US-10-156-761-10049
10	128	18.9	1081	10	US-09-424-951-4
11	127.5	18.8	2150	14	US-10-135-322-17
12	126.5	18.7	227	9	US-09-815-242-1131
13	126	18.6	203	15	US-10-156-761-12532
14	125.5	18.5	227	15	US-10-156-761-9975
15	124.5	18.4	232	9	US-09-815-242-13783

16	124	18.3	971	14	US-10-116-048-2	Sequence 2, Appli
17	124	18.3	2471	14	US-10-116-048-4	Sequence 4, Appli
18	123.5	18.2	185	11	US-09-769-787-54	Sequence 54, Appli
19	123	18.1	1447	15	US-10-156-761-8624	Sequence 8624, Ap
20	122.5	18.1	248	15	US-10-156-761-12602	Sequence 12602, A
21	122.5	18.1	1044	15	US-10-101-464A-956	Sequence 956, App
22	122	18.0	248	9	US-09-815-242-12098	Sequence 12098, A
23	121.5	17.9	218	15	US-10-156-761-13746	Sequence 13746, A
24	120.5	17.8	222	9	US-09-815-242-11629	Sequence 11629, A
25	118.5	17.5	232	9	US-09-815-242-10420	Sequence 10420, A
26	118	17.4	890	15	US-10-101-464A-958	Sequence 958, App
27	118	17.4	1270	15	US-10-101-464A-979	Sequence 979, App
28	117.5	17.3	1173	14	US-10-135-322-19	Sequence 19, Appli
29	117.5	17.3	1383	15	US-09-918-508-2	Sequence 2, Appli
30	117	17.3	1383	15	US-10-156-761-13096	Sequence 13096, A
31	115	17.0	762	15	US-10-101-464A-114	Sequence 114, App
32	114.5	16.9	243	15	US-10-156-761-11581	Sequence 11581, A
33	113	16.7	747	12	US-10-100-294A-27	Sequence 27, Appli
34	113	16.5	228	15	US-10-156-761-10048	Sequence 10048, A
35	111.5	16.4	235	10	US-09-738-626-6370	Sequence 6370, Ap
36	111	16.4	669	11	US-09-819-142-24	Sequence 24, Appli
37	111	16.4	1036	10	US-09-918-508-4	Sequence 4, Appli
38	110.5	16.3	104	14	US-10-135-322-30	Sequence 30, Appli
39	110	16.2	573	11	US-09-819-142-10	Sequence 10, Appli
40	110	16.2	599	15	US-10-101-464A-117	Sequence 117, App
41	110	16.2	816	15	US-10-101-464A-827	Sequence 827, App
42	110	16.2	1240	15	US-10-101-464A-976	Sequence 976, App
43	109.5	16.2	100	14	US-10-135-322-29	Sequence 29, Appli
44	109.5	16.2	1092	14	US-10-135-322-18	Sequence 18, Appli
45	107	15.8	197	15	US-10-156-761-10458	Sequence 10458, A

ALIGNMENTS

RESULT 1
US-10-101-464A-127
Sequence 127, Application US/10101464A
Publication No. US20030046728A1
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
FILE REFERENCE: 11000.1020C2
CURRENT APPLICATION NUMBER: US/101.464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
PRIOR FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq For Windows Version 4.0
SEQ ID NO 127

LENGTH: 261

TYPE: PRT

ORGANISM: Pinus radiata

US-10-101-464A-127

Query Match 29.8%; Score 202; DB 15; Length 261;
Best Local Similarity 36.5%; Pred. No. 4.4e-14;
Matches 42; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

QY 18 VLIVDDPLNLIIHKIKKIGGSSQANNGEAVIIHRDGGSSFDLILMDKEMPERDGV 77

Db 137 ILIVDDPLNLIIHKIKKIGGSSQANNGEAVIIHRDGGSSFDLILMDKEMPERDGV 195

QY 78 STTKKLEMEVKSMTIVGVTSLADNEERAPFMEAGLNHCLAKPLTKDKIPLINQ 132

Db 196 EATRLRSRGVKTPIVALTA-NTLQSDKDLFFAGVDFQSKPLSRDLRLVQLLQ 249

RESULT 2

US-09-918-508-8
; Sequence 8, Application US/09918508
; Patent No. US20020177162A1
; GENERAL INFORMATION:
; APPLICANT: KAKIMOTO, TATSUO
; APPLICANT: HIGUCHI, MASAYUKI
; APPLICANT: INOUE, TSUTOMU
; TITLE OF INVENTION: ANALYSIS OF AGONIST-ACTIVITY AND ANTAGONIST-ACTIVITY
; TITLE OF INVENTION: TO CYTOKININ RECEPTOR
; FILE REFERENCE: Q65478
; CURRENT APPLICATION NUMBER: US/09/918,508
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: JP 2001-073812
; PRIOR FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 8
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-918-508-8

Query Match 25.3%; Score 171.5; DB 10; Length 118;
Best Local Similarity 33.9%; Pred. No. 3.2e-11;
Matches 37; Conservative 32; Mismatches 37; Indels 3; Gaps 2;
QY 18 VLIVDDPLNLIHKKIKAIGISQATNGGEAVIIHRDGGSSFDLIMDKEMPERDGV 77
Db 6 ILVVDHPIFNRLLAQLGSLGVQCKTANGVDALNVL--SKNHIDIVLSDVNMFMNDGY 63
QY 78 STTKLREMEVKSMIVGVTSIADNEERAFWEAGLNHCLAKPLTKDKI 126
Db 64 RLQRIQLGLTLPVIGVTANALAEKQRC-LESQMDSCLSKPEVTLVVI 111

RESULT 3

US-09-801-368-328
; Sequence 328, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 328
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-328

Query Match 22.1%; Score 150; DB 10; Length 622;
Best Local Similarity 28.1%; Pred. No. 7.3e-08;
Matches 38; Conservative 32; Mismatches 45; Indels 20; Gaps 4;
QY 12 IKKLNVLIVDDPLNLIHKKIKAIGISQATNGGEAVII--HRDGGSSFDLIMD 68
Db 373 URKGFHVLIVDDDAVSQQLCKFKRYGCTVVVSDGLSAISTLEKYR-----YDIVLMD 427
QY 69 KEMPERDGVSTTKLREMEVKSMIVGVTSIADNEERAFWEAGLNHCLAKPLTKDKI-- 126
Db 428 IWPNLDCATATSVRSFNETPIIATMTGNIMN-QDLITYLQHGMDILAKPFRDLDHS 486
QY 127 -----IPLNQ 132
Db 487 ILIRYLKDRIFLCEQ 501

RESULT 4

US-101-101-464A-909
; Sequence 909, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 909
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-101-101-464A-909

Query Match 21.5%; Score 146; DB 15; Length 1018;
Best Local Similarity 25.2%; Pred. No. 4e-07;
Matches 41; Conservative 35; Mismatches 53; Indels 34; Gaps 3;
QY 3 TKSMGDIEKIKKLNVLIVDDPLNLIHKKIKAIGISQATNGGEAVIIHRDGGSSP 62
Db 858 SSSIIVEEPKPKPKILLVDNKNVMTQSMKQLGHSMGVNNGVEAV--HAVQOOSTY 915
QY 63 DLILMDKEMPERDGVSTTKKL-----REMEYKSM 91
Db 916 DLILMDVCMVYNGVQATRIIRSFEEENGNDAAVNAVIGIELVSSDLSNCHSSRESKERV 975
QY 92 IGVVTSIADNEERAFWEAGLNHCLAKPLTKDKIIFLINQLM 134
Db 976 IIAMTANALSESADCFAN-GMDSFVSKPVTQKLCLEQYL 1017

RESULT 5

US-09-801-368-332
; Sequence 332, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug

```

; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 332
; LENGTH: 1220
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-332

Query Match      21.4%; Score 145; DB 10; Length 1220;
Best Local Similarity 29.0%; Pred. No. 6.6e-07;
Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;

QY      8 DIEKKKLVNVLVDDPLNLIHKKIKA---IGGIS--QTANNGBEA---VHHRDGG 59
DB      1080 DDKNKTSVKVLWEDNVN---QEVKRVNLEGIENIELACDQGEAFDKVEIKTSKG 1135
QY      60 SSFDLILDKEMPERDGVSTTKKL-RENEVKSMIVGVTSIADNEEERRRAMEAGLNHCLA 118
DB      1136 ENYNIEMDVQMPKVDGLLSTKMTRRDLGYTSPIVALTAADDNSNIKEC-JESGMNGFLS 1194
QY      119 KPLTKDKIIPINQMDA 136
DB      1195 KPIKRPKLKILTEPCAA 1212

RESULT 6
US-09-801-368-366
; Sequence 366, Application US/09801368
; Patent No. US20020128250Al
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 366
; LENGTH: 712
; TYPE: PRT

```

```

; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-366

Query Match      21.2%; Score 143.5; DB 10; Length 712;
Best Local Similarity 27.9%; Pred. No. 4.5e-07;
Matches 39; Conservative 28; Mismatches 42; Indels 31; Gaps 3;

QY      10 EKIKKKLVNVLVDDPLNLIHKKIKAIGGISQTANNGBEAVIHRDGGSFDLILMDK 69
DB      498 EKVPFKINVLIVEDNVNQILGSLFKHKHSYKLAHQGEAVNIWKEGG--LHLIFMDL 555
QY      70 EMPERDGVSTTKKLREMEVKSS-----MIVGVTSIADN 101
DB      556 QLPVLGIEAAKQIRDFEKQNGIGIQKSLNNSHNSLEKTSKFPSCAPVILVATASNSQ 615
QY      102 EEERRAPMEAGLNHCLA 121
DB      616 MDKEKALL-SGCNDYLTTPV 634

RESULT 7
US-10-126-120-2
; Sequence 2, Application US/10126120
; Publication No. US20030108526A1
; GENERAL INFORMATION:
; APPLICANT: Takei, Kentaro
; APPLICANT: Sakakibara, Hitoshi
; TITLE OF INVENTION: MICROORGANISMS FOR USE IN THE MEASUREMENT OF ENVIRONMENTAL FACTORS
; FILE REFERENCE: 11127-004001
; CURRENT APPLICATION NUMBER: US/10/126,120
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: JP 2001-291059
; PRIOR FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Zea mays
US-10-126-120-2

Query Match      20.3%; Score 137.5; DB 15; Length 974;
Best Local Similarity 24.6%; Pred. No. 3.2e-06;
Matches 35; Conservative 32; Mismatches 46; Indels 29; Gaps 3;

QY      17 NVLIVDDPLNLIHKKIKAIGGISQTANNGBEAVI---HRDGGSFDLILMDKEMP 72
DB      830 NILVVDNKNLKVAAALKKYGANVSCVSGKDAISLLQPPHR-----FDACFMDVQMP 884
QY      73 ERDGVSTTKKLREMEVKSMIVGVTSIADNEEERRA-----FMEAG 112
DB      885 EMDGFEATGQIQMELKANBERKNKLASIEGSTTBYHLPVLAMTADVTQATYECIKSG 944
QY      113 LNHCLAKPLTKDKIIPINQML 134
DB      945 MDGVVSKFPDEEQLYQAVSRV 966

RESULT 8
US-09-918-508-7
; Sequence 7, Application US/09918508
; Patent No. US20020177162A1
; GENERAL INFORMATION:
; APPLICANT: KAKIMOTO, TATSUO
; APPLICANT: HIGUCHI, MASAYUKI
; APPLICANT: INOUE, TSUTOMU
; TITLE OF INVENTION: ANALYSIS OF AGONIST-ACTIVITY AND ANTAGONIST-ACTIVITY
; TITLE OF INVENTION: TO CYTOKININ RECEPTOR
; FILE REFERENCE: Q65478
; CURRENT APPLICATION NUMBER: US/09/918,508
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: JP 2001-073812
; PRIOR FILING DATE: 2001-03-15

```

```
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-918-508-7

Query Match
Best Local Similarity 20.2%; Score 137; DB 10; Length 125;
Matches 37; Conservative 38; Mismatches 37; Indels 14; Gaps 6;

QY 16 LNVLIYDDPPLNLIHEKIKA---IGGTS--QTANNGEAA---VIIHRDGGSSFDLILM 67
Db 4 VKILVVEDNHN---QEVKRLNLEGHENIELACDQGEAFDKVKELTSGENTNMFIM 59

QY 68 DKEMPERDGVSTTKGL-REMEVKSMIVGTVSLADNEEERAFMEAGLNHCLAKPLTKDKI 126
Db 60 DVQMPKVDGLLSTKMIRDLGTVSPVALTAFADDSNIKEC-LESCMNGFLSKPIKRPKL 118

QY 127 IPLINQ 132
Db 119 KTIUTE 124

RESULT 9
US-10-156-761-10049
; Sequence 10049, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10049
; LENGTH: 1829
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10049

Query Match
Best Local Similarity 19.5%; Score 132; DB 15; Length 1829;
Matches 39; Conservative 24; Mismatches 51; Indels 10; Gaps 4;

QY 18 VLIYDDPPLNLIHEKIKAIGGISQTANNGEAAVII---HRDGGSSFDLILMDKEMPER 74
Db 1710 VLIYDDIRNVFALTSVLEQHGSLVYAENGRGIEVLQCHDD-----VTVVLDMDIMDEM 1765

QY 75 DGVSTTKKLREME--VKSMIVGTVSLADNEEERAFMEAGLNHCLAKPLTKDKIPLINQ 132
Db 1766 DGVATTTAIRMPQFAGLPPIALTAKAMKGDREKA-IESGASDVYTKPVDPPHLLTVMEQ 1824

QY 133 LMDA 136
Db 1825 WURA 1828

RESULT 10
US-09-424-951-4
; Sequence 4, Application US/09424951
; Patent No. US20020137034A1
```

```
; GENERAL INFORMATION:
; APPLICANT: SRIKANTHA, THYAGARAJAN
; APPLICANT: SOLL, DAVID R.
; TITLE OF INVENTION: CANDIDA ALBICANS TWO-COMPONENT HYBRID KINASE GENE,
; FILE REFERENCE: 087714/0113, AND USE THEREOF
; CURRENT APPLICATION NUMBER: US/09/424,951
; CURRENT FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: PCT/US98/11658
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048,914
; PRIOR FILING DATE: 1997-06-06
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-424-951-4

Query Match
Best Local Similarity 18.9%; Score 128; DB 10; Length 1081;
Matches 43; Conservative 31; Mismatches 48; Indels 30; Gaps 7;

QY 5 SMGDIEKIKKLNVLIVDDPPLNLIHEKIKAIGGISQTANNGEAAV-IHRDGGSSFD 63
Db 865 SQNSDESRYK--ILLAEDNLVNQKLAVRILEKQHLVEVNGLEAYEAIKR---NKYD 919

QY 64 LILMDKEMPERDGVSTTKKLREMEVKS-----MIVGVTS---LADNEEERAFME 110
Db 920 VVIMDVQMPVGMGGFEATEKIQWEKKSNDLSLTFRTPIALTAAHMLGDRKS----LA 975

QY 111 AGLNHCCLAKPL-----TKDKIPLINQ 135
Db 976 KGWDVVSFKPLKPKLLMOTIKKCIHNINQLKE 1007

RESULT 11
US-10-135-322-17
; Sequence 17, Application US/10135322
; Publication No. US20020173017A1
; GENERAL INFORMATION:
; APPLICANT: BENFEY, PN
; APPLICANT: HELARIUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AWM
; APPLICANT: KAUPPINEN, L
; APPLICANT: RIIKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17

Query Match
Best Local Similarity 18.8%; Score 127.5; DB 14; Length 2150;
Matches 37; Conservative 20; Mismatches 50; Indels 7; Gaps 3;

QY 13 KKKLNVLIVDDPPLNLIHEKIKAIGGISQTANNGEAAV-IHRDGGSSFDLILMDKEMP 72
Db 2024 RKK--ALIVEDNELNRKVLQKLFKIDWTISFAENGREA-LKEITGRCFDIVFMDCOMP 2080

QY 73 ERDGVSTTK-----KLREMEVKSIVGTVSLADNEEERAFMEAGLNHCLAKPLT 122
Db 2081 VLDGFQTTKIIRSKERNWKEMNIVALSGSSSFVQDCLDSDGMDSFMGKPI 2134
```

RESULT 12

US-09-815-242-11131
; Sequence 11131, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11131
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11131

Query Match 18.7%; Score 126.5; DB 9; Length 227;
Best Local Similarity 29.1%; Pred. No. 6.7e-06;
Matches 34; Conservative 30; Mismatches 46; Indels 5; Gaps 2;

QY 18 VLIVDDPLNLIHEKIIKAIGISQTANNNGEAVIIHRDGGSSFDLILMDKEMPERDGV 77
Db 4 LLLVDDIIELTSTLLEGGFDVETANNGLAEL---QKLNESYKVLDDVMPKPLNGI 60

QY 78 STTKKLEMEVKSVMIVGVTSADNNEERAFMEAGLNHCLAKPLTKDKKIPLINQLM 134
Db 61 ETLKEIR--KVSNPVWMLTARGEDIDRVIGLELGGADDCPLKPFNDRELIARIKAIL 115

RESULT 13

US-10-156-761-12532
; Sequence 12532, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-29

Query Match 18.7%; Score 125.5; DB 15; Length 227;
Best Local Similarity 29.2%; Pred. No. 8.6e-06;
Matches 38; Conservative 28; Mismatches 53; Indels 11; Gaps 5;

QY 10 EKIKKKLVNLIIVDDPLNLIHEKIIKAIG---GIS--QTANNNGEAVIIHRDGGSSFDL 64
Db 4 EBSRKPAAVWVADD---QTVVREGIVMLLGLPGIEVVGAGDGHEAVKLVAE--LNPDV 58

QY 65 ILMDKEMPERDGVSTTKKLEMEVKSVMIVGVTSADNNEERAFMEAGLNHCLAKPLTKD 124
Db 59 VLMDLRMPRCDDVETATRRIRAEHPGTQVVVLTITVADDESIFPA--LRAGARGYITKAGGD 117

QY 125 KIPLINQLM 134
Db 118 EIVRAVESVL 127

RESULT 15

US-09-815-242-13783
; Sequence 13783, Application US/09815242

; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12532
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12532

Query Match 18.6%; Score 126; DB 15; Length 203;
Best Local Similarity 29.5%; Pred. No. 6.5e-06;
Matches 36; Conservative 20; Mismatches 46; Indels 20; Gaps 3;

QY 17 NVLIVDDPLNLIHEKIIKAIGISQ--TANNNGEAVIIHRDGGSSFDLILMDKEMPER 74
Db 3 SYLVCDSDSFLAREALRAVATVPGVERTTAANGEE--VLRRWGAQRSDDLILMDVMPGL 60

QY 75 DGVSTTKKLEMEVKSVMIVGVTSADNNEERAFMEAGLNHCL 118
Db 61 GGVETVRKLLSADPGARIIMLTVAEDLDGVALAVAGARGYLHKDASRAELRATVTTQALA 120

QY 119 KP 120
Db 121 DP 122

RESULT 14

US-10-156-761-9975
; Sequence 9975, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9975
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9975

Query Match 18.5%; Score 125.5; DB 15; Length 227;
Best Local Similarity 29.2%; Pred. No. 8.6e-06;
Matches 38; Conservative 28; Mismatches 53; Indels 11; Gaps 5;

QY 10 EKIKKKLVNLIIVDDPLNLIHEKIIKAIG---GIS--QTANNNGEAVIIHRDGGSSFDL 64
Db 4 EBSRKPAAVWVADD---QTVVREGIVMLLGLPGIEVVGAGDGHEAVKLVAE--LNPDV 58

QY 65 ILMDKEMPERDGVSTTKKLEMEVKSVMIVGVTSADNNEERAFMEAGLNHCLAKPLTKD 124
Db 59 VLMDLRMPRCDDVETATRRIRAEHPGTQVVVLTITVADDESIFPA--LRAGARGYITKAGGD 117

QY 125 KIPLINQLM 134
Db 118 EIVRAVESVL 127

RESULT 15

US-09-815-242-13783
; Sequence 13783, Application US/09815242

```

; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13783
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13783

```

```

Query Match      18.4%; Score 124.5; DB 9; Length 232;
Best Local Similarity 28.2%; Pred. No. 1.1e-05;
Matches 33; Conservative 31; Mismatches 48; Indels 5; Gaps 2;

Qy      18 VLIYDDPLNLIIEHKIKRIGGISTANNGEEAVIHRDGGSSFDLILMDKEMPERDGV 77
Db      4 ILLVDDRELTSLKELLEMEGFNVLVAHDGEQALELDD---SIDLLLDVWMEKNGI 60

Qy      78 STTKKLEMEVKSINIVGVTSLADNEERRAFMEAGLNHCLAKPLTKDKIPLINOLM 134
Db      61 DTLKALRQTHQTPVI--MLTARGNELDRVLGLELGADDYLPKPFNDRELVARAIL 115

```

Search completed: August 14, 2003, 16:59:18
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 29 Seconds
(without alignments)

198.423 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678
Sequence: 1 MATKSMGDIKKKLNVL.....LAKPLTKDKIPLINQLMDA 136

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A.COMB.pdp:*
2: /cgn2_6/prodata/1/aa/5B.COMB.pdp:*
3: /cgn2_6/prodata/1/aa/5A.COMB.pdp:*
4: /cgn2_6/prodata/1/aa/5B.COMB.pdp:*
5: /cgn2_6/prodata/1/aa/PCTUS.COMB.pdp:*
6: /cgn2_6/prodata/1/aa/backfiles.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	202	29.8	261	4	US-09-228-986-127
2	152.5	22.5	946	4	US-09-328-352-7973
3	150	22.1	1014	4	US-09-252-991A-3138
4	148	21.8	947	4	US-09-252-991A-29359
5	147	21.7	1627	4	US-09-252-991A-20395
6	145	21.4	1220	2	US-08-843-530B-36
7	144	21.2	162	4	US-09-252-991A-17557
8	139.5	20.6	1281	2	US-08-843-530B-6
9	139.5	20.6	1298	2	US-08-843-530B-2
10	139.5	20.6	1298	2	US-08-843-530B-4
11	139.5	20.6	1298	2	US-08-843-530B-34
12	139.5	20.6	1441	4	US-09-252-991A-28143
13	138.5	20.4	234	4	US-09-634-238-220
14	135.5	20.0	760	4	US-09-252-991A-25928
15	135.5	20.0	860	4	US-09-252-991A-26112
16	133.5	19.7	918	2	US-08-843-530B-35
17	131	19.3	764	4	US-09-252-991A-18607
18	131	19.3	1036	4	US-09-252-991A-27075
19	130	19.2	1081	2	US-08-843-530B-18
20	130	19.2	1117	2	US-08-843-530B-33
21	127	18.7	506	4	US-09-252-991A-18168
22	125	18.4	227	4	US-09-107-532A-6769
23	124.5	18.4	227	4	US-09-328-352-5134
24	124	18.3	971	3	US-09-112-450-2
25	124	18.3	971	4	US-09-419-291A-2
26	124	18.3	2471	3	US-09-112-450-4
27	124	18.3	2471	4	US-09-419-291A-4

28 122.5 18.1 265 4 US-09-328-352-6073 Sequence 6073, Ap
29 122 18.0 311 4 US-09-252-991A-22932 Sequence 22932, A
30 120.5 17.8 811 4 US-09-252-991A-22216 Sequence 22216, A
31 119 17.6 484 4 US-09-252-991A-31677 Sequence 31677, A
32 117 17.3 320 4 US-09-328-352-6809 Sequence 6809, Ap
33 115.5 17.0 331 4 US-09-252-991A-23765 Sequence 23765, A
34 115 17.0 762 4 US-09-228-986-114 Sequence 114, App
35 113 16.7 302 4 US-09-252-991A-29564 Sequence 29564, A
36 112.5 16.6 129 4 US-09-252-991A-16981 Sequence 16981, A
37 112.5 16.6 659 4 US-09-252-991A-17904 Sequence 17904, A
38 111.5 16.4 212 4 US-09-107-532A-6859 Sequence 6859, Ap
39 111.5 16.4 234 4 US-09-107-532A-6746 Sequence 6746, Ap
40 110.5 16.3 256 4 US-09-328-352-5430 Sequence 5430, Ap
41 110 16.2 599 4 US-09-228-986-117 Sequence 117, App
42 109 16.1 245 4 US-09-134-001C-3779 Sequence 3779, Ap
43 109 16.1 257 4 US-09-328-352-7397 Sequence 7397, Ap
44 108.5 16.0 246 4 US-09-107-532A-5040 Sequence 5040, Ap
45 108.5 16.0 250 4 US-09-634-238-221 Sequence 221, App

ALIGNMENTS

RESULT 1
US-09-228-986-127
; Sequence 127, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 127
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-127

Query Match 29.8%; Score 202; DB 4; Length 261;
Best Local Similarity 36.5%; Pred. No. 7e-16;
Matches 42; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

QY 18 VLIVDDDPNLIIHEKIIKAIGGISQTANNGEAVIHRDGGSSFDLLIMDKEMPERDGV 77
DB 137 ILLVETQINRIIFGRVLQSLNLYCEAENGKAVDYFKQ-GRYIDLVLMDKEMPVMDGH 195
QY 78 STTKKLEMEYKSMIVGVTSIADNEERRAFMEAGLNHCLAKPLTKDKIPLINQ 132
DB 196 EATRQLRSMGVKTFIVALTA-NTLQSKDLFFEAQVDDFQSKPLSRDLVLQDQ 249

RESULT 2

US-09-328-352-7973
; Sequence 7973, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breston et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7973
; LENGTH: 946
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7973


```
Query Match      22.5%; Score 152.5; DB 4; Length 946;
Best Local Similarity 30.2%; Pred. No. 3.4e-09;
Matches 39; Conservative 34; Mismatches 41; Indels 15; Gaps 4;

QY 16 LNVIVDDPDLNLIHKKIKAIGGISTQANNGEEAVIHRDGGSS----FDLILMDKEM 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 674 LHILAVDDHLPNLIVLEALUGELNVKTKALSGQEAALNIQSRIOQKLKPPDLVFMDIQM 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 72 PERDGVSTTKKLEEM-----EVKSMIVGVT--SLADNEEERRAFMEAGLNHCLAKPLTK 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 734 FVMSGIDTTRAIRSLSTLDGEMQLPIIALTAHALAD---EQKLLKVGMDYVTKPIQM 790
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 124 DKIIPLINQ 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 791 EQIIQILITQ 799
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
US-09-252-991A-31338
; Sequence 31338, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31338
; LENGTH: 1014
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31338

Query Match      22.1%; Score 150; DB 4; Length 1014;
Best Local Similarity 25.5%; Pred. No. 7.4e-09;
Matches 35; Conservative 36; Mismatches 56; Indels 10; Gaps 4;

QY 2 ATKSMGDIKIK--KLVNLIVDDPDLNLIHKKIKAIGGISTQANNGEEAVIHRDGG 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 868 ARSGVAEVEERKARALSILLADHPNRLITLMQLESIGHRVSTEDGEAA--FERWQG 925
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 60 SSFDLILMDKEMPERDGVSTTKKLEMEV-----KSMIVGVTSLADNEEERRAFMEAGLN 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 926 EDFDVITDGMFMDGVELAIRRSQEAALGGRRCLVIALTASAEKDALERC-LAAGWD 984
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 115 HCLAKPLTKDKIIPLIN 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 985 RVLFKPTTLDELAFALN 1001
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-09-252-991A-29359
; Sequence 29359, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 29359
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29359

Query Match      21.8%; Score 148; DB 4; Length 947;
Best Local Similarity 28.8%; Pred. No. 1.1e-08;
Matches 34; Conservative 33; Mismatches 43; Indels 8; Gaps 4;

QY 18 VLIIVDDPDLNLIHKKIKAIGGISTQANNGEEAV-IHRDGGSSFDLILMDKEMPERDG 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 689 LLCVDNDFANLLVQTLSDLGAVTAVDVSGYAALEVYQRE---RFDLVFMDVQPMQMDG 745
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 77 VSTTKKLEMEVKSMI--VGVTSLADN--EEERRAFMEAGLNHCLAKPLTKDKIIPLI 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 745 RQATEAIRWEAREVSPVPVIALTAHALSNEKALLQAGNDDYLTKPIDEQQLAQVV 803
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-09-252-991A-20395
; Sequence 20395, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20395
; LENGTH: 1627
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20395

Query Match      21.7%; Score 147; DB 4; Length 1627;
Best Local Similarity 31.9%; Pred. No. 3.3e-08;
Matches 37; Conservative 26; Mismatches 45; Indels 8; Gaps 3;

QY 14 KCLNVLIVDDPDLNLIHKKIKAIGGISTQANNGEEAVIHRDGGSSFDLILMDKEMPE 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1371 RLQLVIVDDHVNQLIHLQQLSFLGHVDEAEAGLGNLNLWH--GQFFDVIITDCHMPL 1428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 74 RDGVSTTKKLEMEVKS-----MIVGVTSLADNEEERRAFMEAGLNHCLAKPLTKD 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1429 MSGSGLARSIRQEREENGEEPPVVIIGLTADAQPEIERC-IOAGNNECLIKPIGLD 1483
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-08-843-530B-36
; Sequence 36, Application US/08843530B
; Patent No. 5939306
; GENERAL INFORMATION:
; APPLICANT: Selitrennikoff, Claude
; APPLICANT: Agnan, Jacqueline
; APPLICANT: Alex, Lisa A.
; APPLICANT: Simon, Melvin I.
; TITLE OF INVENTION: Osmosensing Histidine Kinases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
```

DB 95 EATRLIRREERAQGWPRVPIVTAHILD--EHRRAGIEAGMDAYLGKPFVDRALYATLE 152

QY 132 QLM 134

DB 153 RLL 155

RESULT 8

US-08-843-530B-6

Sequence 6, Application US/08843530B

Patent No. 5939306

GENERAL INFORMATION:

APPLICANT: Selitrennikoff, Claude

APPLICANT: Agnan, Jacqueline

APPLICANT: Alex, Lisa A.

APPLICANT: Simon, Melvin I.

TITLE OF INVENTION: Osmosensing Histidine Kinases

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/843,530B

FILING DATE: 16-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MacKnight, Karin T.

REGISTRATION NUMBER: 38,230

REFERENCE/DOCKET NUMBER: UTC-02717

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1281 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-843-530B-6

Query Match 20.6%; Score 139.5; DB 2; Length 1281;

Best Local Similarity 31.1%; Pred. No. 1.8e-07;

Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

QY 2 ATKSMGDIETKIKKLNVLIVDDPLNLIHEKIIKAIGGISQTANNGEAV-IHHRDGS 60

DB 1075 ATPSLAD--NTKSFELLAEADNTVQNLAVKILEKHYHVTVVVGNGEAVEAVKR---K 1128

QY 61 SPDLILMKEMPERDGVSTTKKUREME-----VKSMIVGVTSIADNEERAFMEAGLNH 115

DB 1129 KFDVILMDVQPMGPGFEATAKIREYERLSGSRTPIALTAFA-MMGDRKCIQAQOMDE 1187

QY 116 CLAKPLTKDKIPLI 130

DB 1188 YLSKPLQONHLIQT 1202

RESULT 9

US-08-843-530B-2

Sequence 2, Application US/08843530B

Patent No. 5939306

GENERAL INFORMATION:

APPLICANT: Selitrennikoff, Claude

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/843,530B

FILING DATE: 16-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MacKnight, Karin T.

REGISTRATION NUMBER: 38,230

REFERENCE/DOCKET NUMBER: UTC-02717

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 1220 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

US-08-843-530B-36

Query Match 21.4%; Score 145; DB 2; Length 1220;

Best Local Similarity 29.0%; Pred. No. 3.8e-08;

Matches 40; Conservative 40; Mismatches 44; Indels 14; Gaps 6;

QY 8 DIEKIKKLNVLIVDDPLNLIHEKIIKA---IGGIS--QTANNGEAA---VLIHRDGG 59

DB 1080 DDDKNTSKVILVVDNHN---QEVIKRLNLEGIEIACDQGAFAKVKELTSKG 1135

QY 60 SSPDLILMKEMPERDGVSTTKKL-REMEVKSMIVGVTSIADNEERAFMEAGLNHCLA 118

DB 1136 ENYMFIMFQVQPKVDGLSLTKWIRDLGYTSPIVALTAFADDSNIKEC-LESNGWGF 1194

QY 119 KPLTKDKIPLINQLMDA 136

DB 1195 KPIKPKLTKILTEFCAA 1212

RESULT 7

US-09-252-991A-17557

Sequence 17557, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 17557

LENGTH: 162

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17557

Query Match 21.2%; Score 144; DB 4; Length 162;

Best Local Similarity 30.9%; Pred. No. 2.6e-09;

Matches 38; Conservative 28; Mismatches 49; Indels 8; Gaps 3;

QY 18 VLIVDDPLNLIHEKIIKAIGGISQTANNGEAVIHRDGSFPLILMKEMPERDGV 77

DB 35 ILVAEDNPVNLVVRGFLAKRGVAVRLAGNRLALDEYLRDENGQLILMDGEMPMDF 94

QY 78 STTKUREME-----VKSMIVGVTSIADNEERAFMEAGLNHCLAKPLTKDKIPLIN 131

APPLICANT: Agnan, Jacqueline
APPLICANT: Alex, Lisa A.
APPLICANT: Simon, Melvin I.
TITLE OF INVENTION: Osmosensing Histidine Kinases
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,530B
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: UTC-02717
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-843-530B-2

Query Match 20.6%; Score 139.5; DB 2; Length 1298;
Best Local Similarity 31.1%; Pred. No. 1.8e-07;
Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

QY 2 ATKSMGDIKIKKLNVLIVDDPLNLIHKKIKAIGGISQTANNGBEAV-IHHRDGS 60
DB 1075 ATPSLAD---NTKSFELLAEADNTVNQRLAVKILEKHYHVVTVVVGNGEEAVEAVKR--K 1128

QY 61 SFDLIIMDKEMPERDGVSTTKKLEME-----VKSMIVGVTSLADNEERRAFMEAGLNH 115
DB 1129 KFDVILMDVQPMGPGFEATKIREYSLGSRTPILTAHA-MMGDREKCIQAQMD 1187

QY 116 CLAKPLTKDKIPLI 130
DB 1188 YLSKPLQONHLIQT 1202

RESULT 10
US-08-843-530B-4
Sequence 4, Application US/08843530B
Patent No. 5939306
GENERAL INFORMATION:
APPLICANT: Selitrennikoff, Claude
APPLICANT: Agnan, Jacqueline
APPLICANT: Alex, Lisa A.
APPLICANT: Simon, Melvin I.
TITLE OF INVENTION: Osmosensing Histidine Kinases
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,530B
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: UTC-02717
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-843-530B-2

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,530B
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: UTC-02717
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-843-530B-4

Query Match 20.6%; Score 139.5; DB 2; Length 1298;
Best Local Similarity 31.1%; Pred. No. 1.8e-07;
Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

QY 2 ATKSMGDIKIKKLNVLIVDDPLNLIHKKIKAIGGISQTANNGBEAV-IHHRDGS 60
DB 1075 ATPSLAD---NTKSFELLAEADNTVNQRLAVKILEKHYHVVTVVVGNGEEAVEAVKR--K 1128

QY 61 SFDLIIMDKEMPERDGVSTTKKLEME-----VKSMIVGVTSLADNEERRAFMEAGLNH 115
DB 1129 KFDVILMDVQPMGPGFEATKIREYSLGSRTPILTAHA-MMGDREKCIQAQMD 1187

QY 116 CLAKPLTKDKIPLI 130
DB 1188 YLSKPLQONHLIQT 1202

RESULT 11
US-08-843-530B-4
Sequence 34, Application US/08843530B
Patent No. 5939306
GENERAL INFORMATION:
APPLICANT: Selitrennikoff, Claude
APPLICANT: Agnan, Jacqueline
APPLICANT: Alex, Lisa A.
APPLICANT: Simon, Melvin I.
TITLE OF INVENTION: Osmosensing Histidine Kinases
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,530B
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: UTC-02717
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-843-530B-4

```

; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-843-530B-34

Query Match      20.6%; Score 139.5; DB 2; Length 1298;
Best Local Similarity 31.1%; Pred. No. 1.8e-07;
Matches 42; Conservative 28; Mismatches 52; Indels 13; Gaps 5;

QY 2 ATKSMGDIKIKKLVNLIHEDDPLNLIHKKIIGGISTANNGEEAVIHRDGGSG 60
DB 1075 ATPSLAD---NTKSPFLLAEDNTNQLAVKILEKHYHVVTVVNGSEAVEAYKR--K 1128
QY 61 SFLIIMDKEMPERDQVSTTKKLREME-----VKSMIVGVTSADNEERRAFMEAGLNH 115
DB 1129 KFDVILMDVQMPIMGFTEATKIREVERSLGSGRTPIIALTAHA-MWGDREKCIQAWDE 1187
QY 116 CLAKPLTKKIPII 130
DB 1188 YLSKPLQONHLIQT 1202

RESULT 12
US-09-252-991A-28143
; Sequence 28143, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28143
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-28143

Query Match      20.6%; Score 139.5; DB 4; Length 1441;
Best Local Similarity 28.8%; Pred. No. 2.2e-07;
Matches 34; Conservative 30; Mismatches 43; Indels 11; Gaps 3;

QY 16 LNVLIIVDDPPLNLIHKKIIGGISTANNGEEAVIHRDGGSGFDLILMDKEMPERD 75
DB 1182 VRLIIVEDNALNRQVAELLSSEGARVALADGLAGVQVLEASVPFDVLMQMPDID 1241
QY 76 GVSTTKLRE-----NEVKSVMIVGVTSADNEERRAFMEAGLNHCLAKPLTKKII 127
DB 1242 GLEATRRIRADGRFAGFILAMTAN-ASLAD----REACLAAGMNDHVAKPIDKERLV 1294

RESULT 13
US-09-634-238-220
; Sequence 220, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James

```

```

; APPLICANT: Christenson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.1043UI
; CURRENT APPLICATION NUMBER: US/09/634,238
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 220
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
; US-09-634-238-220

Query Match      20.4%; Score 138.5; DB 4; Length 234;
Best Local Similarity 30.2%; Pred. No. 2e-08;
Matches 35; Conservative 31; Mismatches 47; Indels 3; Gaps 2;

QY 18 VLIIVDDPPLNLIHKKIIGGISTANNGEEAVIHRDGGSGFDLILMDKEMPERDGV 77
DB 4 ILIVDEPAITLTLQYNLEAEHYQVETADQGEA--LDKVRSEPFDFIILDLMLPSLSGL 61
QY 78 STTKKLEMEVKSVMIVGVTSADNEERRAFMEAGLNHCLAKPLTKDKIPIINQL 133
DB 62 DVTCKIREKIQTPIMLTA-KDNETKIVGLGADDDVYTKPSPREIARIKAI 116

RESULT 14
US-09-252-991A-25928
; Sequence 25928, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25928
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25928

Query Match      20.0%; Score 135.5; DB 4; Length 760;
Best Local Similarity 28.0%; Pred. No. 2.5e-07;
Matches 35; Conservative 31; Mismatches 54; Indels 5; Gaps 3;

QY 4 KSMGDIKIKKLVNLIHKKIIGGISTANNGEEAVIHRDGGSGFD 63
DB 612 QAAGDAAPVAAGQEILLVEDNPNVOTVIEMLRSLSGYVTLVADGIQAV--RSAERQRYD 669
QY 64 LILMDKEMPERDQVSTTKKLREMEVKSVMIVGVTSADN--EEERRAFMEAGLNHCLAKPL 121
DB 670 AILMDCRLPVLGDYSATREIRAQ-NGRQVPIIALTANALQDRENCIQAQMNDYLAKPF 728
QY 122 TKDKI 126
DB 729 KRAEL 733

RESULT 15
US-09-252-991A-26112
; Sequence 26112, Application US/09252991A

```

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26112
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26112

Query Match 20.0%; Score 135.5; DB 4; Length 860;
Best Local Similarity 26.5%; Pred. No. 3e-07;
Matches 36; Conservative 35; Mismatches 56; Indels 9; Gaps 3;
Qy 1 MATKSMGDIKKKLNVLIVDDPLNLIHKKIKAIGGISCTANNGEEAVIIHRDGG 60
Db 722 IAFSSADTTGEGNTRVLVEDPNVQLVAKGLLHKLGCQVWIAEHGLNALKMLEE--H 779
Qy 61 SFDLILMDKEMPERDGVSTTKKLRM----EVKSMIVGVTSLADNEEEERRAFMEAGLNHC 116
Db 780 PIDLVLMDCNNPVMGDGYEATRQIRDSGRWGGLPIIALTANALPDERERCRA---AGMDY 836
Qy 117 LAKPLTKDKLIPLINQ 132
Db 837 LAKPFHDEKAILDR 852

Search completed: August 14, 2003, 16:51:53
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 16:46:47 ; Search time 83 Seconds
(without alignments)
260.082 Million cell updates/sec

Title: US-09-646-679-15

Perfect score: 678
Sequence: 1 MATKMGDIKKIKLNLVI.....LAKPLTKDKIPLINQLMDA 136

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	678	100.0	136	20	AA1980.DAT.*
2	678	100.0	136	20	AA1981.DAT.*
3	538	79.4	142	21	AA1982.DAT.*
4	538	79.4	142	21	AA1983.DAT.*
5	515	76.0	142	21	AA1984.DAT.*
6	472	69.6	116	20	AA1985.DAT.*
7	437	64.5	104	21	AA1986.DAT.*
8	437	64.5	104	21	AA1987.DAT.*
9	301	44.4	70	21	AA1988.DAT.*

10	280	41.3	66	21	AAG21090	Arabidopsis thalia
11	213	31.4	71	21	AAG21088	Arabidopsis thalia
12	192.5	29.8	261	21	AA1985159	Pinus radiata cell
13	192.5	28.4	67	20	AA1985159	Brassica napus D22
14	171.5	25.3	118	23	ABG70785	E. coli RscC recei
15	156.5	23.1	471	24	AAO26983	Pseudomonas aerugi
16	156.5	23.1	496	24	AAO26985	Pseudomonas aerugi
17	156.5	23.1	666	24	AAO26994	Pseudomonas aerugi
18	156.5	23.1	709	24	AAO26993	Pseudomonas aerugi
19	156.5	23.1	719	24	AAO26992	Pseudomonas aerugi
20	156.5	23.1	802	24	AAO26991	Pseudomonas aerugi
21	156.5	23.1	870	24	AAO26990	Pseudomonas aerugi
22	156.5	23.1	874	24	AAO26989	Pseudomonas aerugi
23	156.5	23.1	887	24	AAO26988	Pseudomonas aerugi
24	156.5	23.1	900	24	AAO26987	Pseudomonas aerugi
25	156.5	23.1	906	24	AAO26986	Pseudomonas aerugi
26	156.5	23.1	931	24	AAO26985	Pseudomonas aerugi
27	155	22.9	922	22	AA1983276	Histidine protein k
28	149	22.0	1373	23	ABP73292	Candida albicans e
29	148.5	21.9	922	22	AA1983274	Histidine protein k
30	148.5	21.9	922	22	AA1983275	Histidine protein k
31	148.5	21.9	950	23	AB193422	Herbicidally activ
32	146	21.5	1018	21	AA1935534	Eucalyptus grandis
33	141	20.8	1146	20	AA1986007	Histidine kinase C
34	139.5	20.6	1281	20	AA198484	osl1 histidine kina
35	139.5	20.6	1298	19	AA198522	N. crassa osp pro
36	139.5	20.6	1298	20	AA198481	osl1p amino acid s
37	139.5	20.6	1298	20	AA198483	osl1 histidine kina
38	137	20.2	125	23	ABG70784	Yeast SLN1 receive
39	134.5	19.8	1122	23	AB192150	Herbicidally activ
40	134	19.8	120	22	AA1986490	Putative P. abyssi
41	133.5	19.7	1023	22	AB198668	Novel human diago
42	131	19.3	136	23	AB191215	Human kinase-like
43	131	19.3	674	21	AA191922	Candida albicans S
44	130	19.2	1081	20	AA198482	osl1 histidine ki
45	130	19.2	1081	20	AA1981600	Candida albicans C

ALIGNMENTS

RESULT 1

AA198483
ID AAY42637 standard; Protein; 136 AA.

AC AAY42637;

XX 10-JAN-2000 (first entry)

DT XX Brassica response regulator protein D22.

DE XX Signal transduction protein; dehiscence; male sterile plant; D22 gene;
KW XX Shatter resistance; oilseed rape; response regulator protein.

OS Brassica napus.

XX WO9949046-A1.

XX 30-SEP-1999.

XX 22-MAR-1999; 99WO-GB00905.

XX 20-MAR-1998; 98GB-0006113.

XX (BIOG-) BIOGENMA UK LTD.

XX Wyatt P, Roberts JA, Whitelaw C;

XX WPI; 1999-580449/49.

XX N-PSDB; AA22974.

PT A nucleic acid encoding a signal transduction protein involved in plant dehiscence, useful for producing shatter resistant male sterile plants

PT -
 XX Claim 4; Fig 1; 71pp; English.
 PS
 XX The invention provides a nucleic acid encoding a signal transduction
 CC protein involved in the process of dehiscence. The nucleic acids and
 CC proteins are useful for regulating or controlling dehiscence of a pod or
 CC an anther in a plant, useful in the production of male sterile plants.
 CC The methods, etc. may be used in the production of shatter resistance or
 CC shatter-delayed plants such as oilseed rape (Brassica napus). The present
 CC sequence represents a B. napus response regulator protein D22.
 XX
 SQ Sequence 136 AA;
 Query Match 100.0%; Score 678; DB 20; Length 136;
 Best Local Similarity 100.0%; Pred. No. 1e-67;
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKSMGDIKKKKLNVLIVDDPLNLIHKKIKAIGGISQTANNGEAVIIHRDGG 60
 DB 1 MATKSMGDIKKKKLNVLIVDDPLNLIHKKIKAIGGISQTANNGEAVIIHRDGG 60
 QY 61 SFDLILMDKEMPERDGVSTTKKLEMEVKSMIVGVTSLADNEEERAFWEAGLNHCLAKP 120
 DB 61 SFDLILMDKEMPERDGVSTTKKLEMEVKSMIVGVTSLADNEEERAFWEAGLNHCLAKP 120
 QY 121 LTKDKIIPILINQLMDA 136
 DB 121 LTKDKIIPILINQLMDA 136
 RESULT 2
 AAY42653
 ID AAY42653 standard; Protein; 136 AA.
 XX
 AC AAY42653;
 XX
 DT 10-JAN-2000 (first entry)
 XX
 DE Brassica response regulator protein D22 putative peptide sequence.
 KW Signal transduction protein; dehiscence; male sterile plant; D22 gene;
 KW shatter resistance; oilseed rape; response regulator protein.
 XX
 OS Brassica napus.
 XX
 PN WO9949046-A1.
 XX
 PD 30-SEP-1999.
 XX
 PF 22-MAR-1999; 98WO-GB00905.
 XX
 PR 20-MAR-1998; 98GB-0006113.
 XX
 PA (BIOG-) BIOGEMMA UK LTD.
 PI Wyatt P, Roberts JA, Whitelaw C;
 XX
 DR WPI; 1999-580449/49.
 DR N-PSDB; AAZ22977.
 XX
 PT A nucleic acid encoding a signal transduction protein involved in plant
 PT dehiscence, useful for producing shatter resistant male sterile plants
 PT -
 XX
 PS Example 2; Fig 6; 71pp; English.
 XX
 CC The invention provides a nucleic acid encoding a signal transduction
 CC protein involved in the process of dehiscence. The nucleic acids and
 CC proteins are useful for regulating or controlling dehiscence of a pod or
 CC an anther in a plant, useful in the production of male sterile plants.
 CC The methods, etc. may be used in the production of shatter resistance or
 CC shatter-delayed plants such as oilseed rape (Brassica napus). The present

CC sequence represents the B. napus D22 putative peptide sequence.
 XX
 SQ Sequence 136 AA;
 Query Match 100.0%; Score 678; DB 20; Length 136;
 Best Local Similarity 100.0%; Pred. No. 1e-67;
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MATKSMGDIKKKKLNVLIVDDPLNLIHKKIKAIGGISQTANNGEAVIIHRDGG 60
 DB 1 MATKSMGDIKKKKLNVLIVDDPLNLIHKKIKAIGGISQTANNGEAVIIHRDGG 60
 QY 61 SFDLILMDKEMPERDGVSTTKKLEMEVKSMIVGVTSLADNEEERAFWEAGLNHCLAKP 120
 DB 61 SFDLILMDKEMPERDGVSTTKKLEMEVKSMIVGVTSLADNEEERAFWEAGLNHCLAKP 120
 QY 121 LTKDKIIPILINQLMDA 136
 DB 121 LTKDKIIPILINQLMDA 136
 RESULT 3
 AAG10549
 ID AAG10549 standard; Protein; 142 AA.
 XX
 AC AAG10549;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 8916.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134378.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140655.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-SEP-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.


```
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.4%; Score 538; DB 21; Length 142;
Best Local Similarity 78.9%; Pred. No. 4.7e-52;
Matches 112; Conservative 8; Mismatches 16; Indels 6; Gaps 2;

QY . 1 MATKSMGDIK-----TKKKL-NVLIVDDPLNLIHKKIKAIGGISOTANNGEAVII 54
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 NATKSTGTEKTKIEVKKKLNVLIVDDPLNRLHEMIKTIKGISOTAKNGEAVIL 60
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 55 HRDGGSPDLIMDKENPERDGVSTTKKREMEVKSMIVGVTSLADNEERRAFMEAGIN 114
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 HRDGEAFDLIMDKENPERDGVSTTKKREMKVTSWIVGTVSVAQESERKAFMEAGIN 120
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 115 HCLAKPLTKDKIIPILNQLMDA 136
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 HCLEKPLTKAKIPFLISHLFDA 142
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
AAG44439
ID AAG44439 standard; Protein; 142 AA.
XX
AC AAG44439;
XX
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55668.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
```


XX SQ Sequence 142 AA;
Query Match 76.08; Score 515; DB 20; Length 142;
Best Local Similarity 76.14; Pred. No. 1.7e-49;
Matches 108; Conservative 8; Mismatches 20; Indels 6; Gaps 2;
QY 1 MATKSMGDIK-----TKKEL-NVLIYDDPLNLIHEKIIKAIGISOTANGSEAVII 54
DB 1 MATKSTGTGTEKTSIEVKKLLNVLIVDDPLNRLHEMIKTIIGISQTAKNGEXVIL 60
QY 55 HRDGGSSFDLILMDKXENPERDGVSTTKKLEMEVKSMIVGVTSIADNNEERAFWEAGLN 114
DB 61 HRDGEAFDILMDKXENPERDGVSTTKKLEMEVKSMIVGVTSIADNNEERAFWEAGLN 120
QY 115 HCLAKPLTKDKIIPILNQLMDA 136
DB 121 HCLEKPLTKAKIFELIHLFDA 142
RESULT 6
AAY42652
ID AAY42652 standard; Protein; 116 AA.
XX
AC AAY42652;
XX
DT 10-JAN-2000 (first entry)
XX
DE Brassica response regulator protein D22B putative peptide sequence.
XX
KW Signal transduction protein; dehiscence; male sterile plant; D22B gene;
XX shatter resistance; oilseed rape; response regulator protein.
XX Brassica napus.
XX
FH Key Location/Qualifiers
FT Misc-difference 2 /note= "unknown"
FT
XX
FN WO9949046-A1.
XX
PN 30-SEP-1999.
XX
PD 22-MAR-1999; 99WO-GB00905.
XX
PR 20-MAR-1998; 98GB-0006113.
XX
PA (BIOG-) BIOGEMMA UK LTD.
XX
PI Wyatt P, Roberts JA, Whitelaw C;
XX
DR WPI; 1999-580449/49.
DR N-PSDB; AAZ22976.
XX
XX A nucleic acid encoding a signal transduction protein involved in plant
dehiscence, useful for producing shatter resistant male sterile plants
-
XX
PS Example 2; Fig 6; 71pp; English.
XX
CC The invention provides a nucleic acid encoding a signal transduction
protein involved in the process of dehiscence. The nucleic acids and
proteins are useful for regulating or controlling dehiscence of a pod or
an anther in a plant, useful in the production of male sterile plants.
CC The methods, etc. may be used in production of shatter resistance or
shatter-delayed plants such as oilseed rape (Brassica napus). The present
sequence represents the B. napus D22B putative peptide sequence.
XX
SQ Sequence 116 AA;
Query Match 69.68; Score 472; DB 20; Length 116;
Best Local Similarity 80.24; Pred. No. 8.4e-45;
Matches 93; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 21 VDDPLNLIHEKIIKAIGISOTANGSEAVIIHRDGGSSFDLILMDKXENPERDGVSTT 80
DB 1 VDDPVIKRLHEIIKSIIGISQTAKNGEAVNHRDGNASFDLILMDKXENPERDGLSAT 60
QY 81 KCLREMEVKSMIVGVTSIADNNEERAFWEAGLNHCLAKPLTKDKIIPILNQLMDA 136
DB 61 KCLREMKVTSIIGVTTIADNNEERAKFWEAGLNHCLAKPLSKAKIIPILNQLMDA 116
RESULT 7
AAG10550
ID AAG10550 standard; Protein; 104 AA.
XX
AC AAG10550;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8917.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.

Db 2 IIKTIGISQAKNGEEAVILHRDGEASFDLILMDKEMPERDGVSTTKLREMKVTSMIV 61

Qy 94 GVTSLADNEEERRRPFWEAGLNHCLAKPTTKDKIPLINOLMDA 136

Db 62 GVTSDADQEEERKAPFWEAGLNHCLAKPTTKAKIPLISHLEDA 104

RESULT 8

AAAG44440
ID :AAG44440 standard; Protein; 104 AA.

AC AAG44440;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55669.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-030439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 23-MAR-1999; 99US-0126264.

PR 23-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 27-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140931.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145224.

PR 23-JUL-1999; 99US-0145226.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 06-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 15-AUG-1999; 99US-0143368.
 PR 17-AUG-1999; 99US-0143175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149502.
 PR 23-AUG-1999; 99US-0149930.
 PR 23-AUG-1999; 99US-0150566.
 PR 25-AUG-1999; 99US-0150884.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 01-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155859.
 PR 24-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 08-OCT-1999; 99US-0157753.
 PR 08-OCT-1999; 99US-0157865.
 PR 08-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 PR 29-OCT-1999; 99US-0162143.

Query Match 64.5%; Score 437; DB 21; Length 104;

Best Local Similarity 84.5%; Pred. No. 5.9e-41;

Matches 87; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 34 IIKAIIGISQTANNGEAAVILHRDGGSSFDLLIMDKEMPERDGVSTTKLRMEVKSIV 93

DB 2 IIKTIGISQTANNGEAAVILHRDGGSSFDLLIMDKEMPERDGVSTTKLRMEVKSIV 61

OY 94 GVTSLADNEERRAFMEAGLNHCLEKPLTKAKIPLIHLFDA 136

Db 62 GVTSLADNEERRAFMEAGLNHCLEKPLTKAKIPLIHLFDA 104

RESULT 9
 AAG21089
 ID AAG21089 standard; Protein; 70 AA.
 XX
 AC AAG21089;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 23518.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EF1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 23-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 08-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 24-JUN-1999;	99US-0140695.	PR 10-SEP-1999;	99US-0153070.
PR 28-JUN-1999;	99US-0140803.	PR 13-SEP-1999;	99US-0153758.
PR 29-JUN-1999;	99US-0140803.	PR 15-SEP-1999;	99US-0154018.
PR 30-JUN-1999;	99US-0141287.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0141842.	PR 20-SEP-1999;	99US-0154779.
PR 01-JUL-1999;	99US-0142154.	PR 22-SEP-1999;	99US-0155139.
PR 02-JUL-1999;	99US-0142055.	PR 23-SEP-1999;	99US-0155486.
PR 06-JUL-1999;	99US-0142390.	PR 24-SEP-1999;	99US-0155659.
PR 08-JUL-1999;	99US-0142803.	PR 28-SEP-1999;	99US-0156458.
PR 09-JUL-1999;	99US-0142920.	PR 29-SEP-1999;	99US-0156596.
PR 12-JUL-1999;	99US-0142977.	PR 04-OCT-1999;	99US-0157117.
PR 13-JUL-1999;	99US-0143542.	PR 05-OCT-1999;	99US-0157753.
PR 14-JUL-1999;	99US-0143624.	PR 07-OCT-1999;	99US-0158029.
PR 15-JUL-1999;	99US-0144005.	PR 08-OCT-1999;	99US-0158232.
PR 16-JUL-1999;	99US-0144086.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144325.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144884.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160880.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 25-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 26-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		

Query March 44.4%; Score 301; DB 21; Length 70;

Best Local Similarity 85.4%; Pred. No. 5.4e-26; Matches 60; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 67 MDKEMPERDGVSTTKKLRNEMVKSMTVGTSLADNBEERRAFMEAGLNHCLAKPLTKDKI 126
 Db 1 MDKEMPERDGVSTTKKLRNEMVKSMTVGTSLADNBEERRAFMEAGLNHCLAKPLTKDKI 60

QY 127 IPLINQLMDA 136
 Db 61 FPLISHLFDA 70

RESULT 10
AAG21090
ID AAG21090 standard; Protein; 66 AA.
XX
XX
AC AAG21090;
XX
DT 17-OCT-2000 (first entry)
XX
DE DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 23519.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 21-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 15-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139762.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 18-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 30-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 41.3%; Score 280; DB 21; Length 66;
Best local similarity 84.8%; Pred. No. 1.1e-23;
Matches 56; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy 71 MPERDGVSTTKKLRENEVKSMIVGVTSLADNEERRAFWEAGLNHCLAKPLTKKIPLI 130
Db 1 MPERDGVSTTKKLRENEVKSMIVGVTSLADNEERRAFWEAGLNHCLAKPLTKKIPLI 60
Qy 131 NQLMDA 136
Db 61 SHLFDA 66

RESULT 11
AAG21088

ID XX AG221088 standard; Protein; 71 AA.
AC AG221088;
XX 17-OCT-2000 (first entry)
DT XX Arabidopsis thaliana protein fragment SEQ ID NO: 23517.
DE XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
KW Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123546.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135639.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

XX DT 27-NOV-2000 (first entry)
 XX DE Pinus radiata cell signalling involved protein SEQ ID NO:127.
 XX KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 XX KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
 XX KW environmental change; development; cell proliferation; differentiation;
 XX KW elongation; survival; disease resistance; nutrient metabolism.
 XX OS Pinus radiata.
 XX PN WO200042171-A1.
 XX PD 20-JUL-2000.
 XX PF 11-JAN-2000; 2000WO-US00724.
 XX PR 12-JAN-1999; 99US-0228986.
 XX PR 01-NOV-1999; 99US-0162866.
 XX PR (GENE-) GENESIS RES & DEV CORP LTD.
 XX PA Strabala TJ, Nieuwenhuizen NJ;
 XX DR WPI; 2000-476052/41.
 XX PT Isolated polynucleotide encoding a polypeptide involved in cell
 XX PT signaling used for generating transgenic plants with modified responses
 XX PT to external signals -
 XX PS Claim 3; Page 104-105; 527pp; English.
 XX CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
 XX CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 XX CC pine (Pinus radiata) also known as Monterey pine). The protein sequences
 XX CC are involved in cell signalling. The polynucleotide and protein
 XX CC sequences can be used to modify the response of plant cells to external
 XX CC signals e.g. environmental changes or pathogens during the growth and
 XX CC development of a plant. They can be used to modify cell proliferation,
 XX CC differentiation, elongation and survival, resistance to disease and
 XX CC nutrient metabolism. Examples of modifications which can be produced are
 XX CC altered fruit ripening and senescence of leaves and flowers e.g. to
 XX CC delay senescence and prolong the life of cut flowers or enhance
 XX CC senescence of reproductive organs to engineer sterile plants. Other
 XX CC modifications can be used to delay senescence in selected cell types or
 XX CC organs providing fruit and vegetables which have a longer shelf life
 XX CC between harvest and consumption, or to decrease branching frequency in
 XX CC forest tree species giving long stretches of valuable knot-free clear
 XX CC wood which can be used in solid timber furniture and veneers.
 XX SQ Sequence 261 AA;
 Query Match 29.8%; Score 202; DB 21; Length 261;
 Best Local Similarity 36.5%; Pred. No. 3.7e-14;
 Matches 42; Conservative 34; Mismatches 37; Indels 2; Gaps 2;
 QY 18 VLIVDDPLNLIHKKIKAIGISQTANNGEEAVIHRDGGSSFDLILMDKEMPREDGV 77
 DB 137 ILLVEDTQINRIIFGRVLQSLNLYCEEAEKGVAVDYFKQ-GRTYDLVLDMDKMPVNDGH 195
 QY 78 STTKLREMEVKSMTIVGVTSLADNEEEERAFVAGLNHCLAKPLTKDKIPLINQ 132
 DB 196 EXTRQLRSGVKTPIVALTA-NTLQSDKDLTFEAGVDDFQSKPLSRDLVQLLDQ 249
 RESULT 13
 AAY42644
 ID AAY42644 standard; Protein; 67 AA.
 XX AC AAY42644;
 XX DT 10-JAN-2000 (first entry)

XX DE Brassica napus DZ2B partial fragment.
 XX KW Signal transduction protein; dehiscence; male sterile plant; DZ2B gene;
 XX KW shatter resistance; oilseed rape; response regulator protein.
 XX OS Brassica napus.
 XX PN WO9949046-A1.
 XX PD 30-SEP-1999.
 XX PF 22-MAR-1999; 99WO-GE00905.
 XX PR 20-MAR-1998; 98GB-0006113.
 XX PA (BIOG-) BIOGENMA UK LTD.
 XX PI Wyatt P, Roberts JA, Whitelaw C;
 XX DR WPI; 1999-580449/49.
 XX DR N-FSDS; AAZ22975.
 XX PT A nucleic acid encoding a signal transduction protein involved in plant
 XX PT dehiscence, useful for producing shatter resistant male sterile plants
 XX PT -
 XX PS Example 2; Fig 5; 71pp; English.
 XX CC The invention provides a nucleic acid encoding a signal transduction
 XX CC protein involved in the process of dehiscence. The nucleic acids and
 XX CC proteins are useful for regulating or controlling dehiscence of a pod or
 XX CC anther in a plant, useful in the production of male sterile plants.
 XX CC The methods, etc. may be used in the production of shatter resistance or
 XX CC shatter-delayed plants such as oilseed rape (Brassica napus). The present
 XX CC sequence represents the partial fragment of B. napus DZ2B protein.
 XX SQ Sequence 67 AA;
 Query Match 28.4%; Score 192.5; DB 20; Length 67;
 Best Local Similarity 65.7%; Pred. No. 6.7e-14;
 Matches 44; Conservative 6; Mismatches 10; Indels 7; Gaps 2;
 QY 1 MAT--KSMGDIK-----IKKLNVLIVDDPLNLIHKKIKAIGISQTANNGEEAVI 53
 DB 1 MATSTGTGDIKTSVEVKKLNVLIVDDTIVTKLHENIIRKSGISQTANNGEEAVN 60
 QY 54 IHRDGS 60
 DB 61 IHRDGA 67
 RESULT 14
 ABG70785
 ID ABG70785 standard; Protein; 118 AA.
 XX AC ABG70785;
 XX DT 09-DEC-2002 (first entry)
 XX DE B. coli RacC receiver region of histidine kinase.
 XX KW RacC; antagonist; agonist; cytokinin receptor;
 XX KW receptor; signal transduction; histidine kinase; hormone; cell division;
 XX KW cell differentiation; agriculture; growth regulator; harvest.
 XX OS Escherichia coli.
 XX PN EP1241182-A2.
 XX PD 18-SEP-2002.
 XX DT 13-MAR-2002; 2002EP-0005749.

